

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 00:24:29 ; Search time 49.62 Seconds  
(without alignments)  
367.112 Million cell updates/sec

Title: US-09-786-130-1

Perfect score: 164  
Sequence: 1 TPVSGTTATATASVSTKDP.....TLPTQELPSHSPQGG 164

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_032802:\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	164	21	AAV83951 Human mature inter
2	164	100.0	192	20	AAW98004 Human interleukin-
3	164	100.0	194	21	AAV96292 Human IGFAM-4 immu
4	164	100.0	194	21	AAV83985 Full length human
5	139	84.8	197	20	AAW98006 Human interleukin-
6	137	83.5	137	21	AAV83982 Human IL-18 bindin
7	97	59.1	161	20	AAW98007 Human interleukin-
8	48	29.3	113	20	AAW98005 Human interleukin-
9	47	28.7	47	21	AAV83984 Human IL-18 bindin
10	40	24.4	40	20	AAW98008 Human interleukin-
11	24	14.6	55	20	AAV12716 Human 5' EST secre

12	20	12.2	117	21	AAV83986	Mouse IL-18 bindin
13	20	12.2	165	21	AAV83952	Mouse mature inter
14	20	12.2	193	21	AAV83989	Full length mouse
15	19	11.6	49	21	AAV83983	Human IL-18 bindin
16	15	9.1	55	20	AAV12536	Human 5' EST secre
17	13	7.9	23	21	AAV83958	Human interleukin
18	12	7.3	29	21	AAV83960	Human interleukin
19	11	6.7	11	21	AAV83955	Human interleukin
20	9	5.5	9	21	AAV83954	Human interleukin
21	8	4.9	8	21	AAV83956	Human interleukin
22	8	4.9	8	21	AAV83957	Human interleukin
23	8	4.9	520	19	AAW79001	Human interleukin
24	8	4.9	520	21	AAW23467	Maize cell death s
25	8	4.9	571	22	AAW27242	zea mays IL51 prot
26	8	4.9	634	22	ABW66245	Human ERMAD-20 SEQ
27	8	4.9	737	22	ABW62878	Drosophila melanog
28	8	4.9	840	20	AAW93570	Novel human diagno
29	8	4.9	840	20	AAW93569	Human conductin pr
30	7	4.3	840	20	AAW93569	Human conductin pr
31	7	4.3	7	21	AAV83962	Human interleukin
32	7	4.3	7	21	AAV83964	Human interleukin
33	7	4.3	7	21	AAV83966	Human interleukin
34	7	4.3	10	21	AAV83979	Human interleukin
35	7	4.3	10	21	AAV83959	Human interleukin
36	7	4.3	54	22	AAW42168	Alpha-L-arabinofur
37	7	4.3	86	22	AAW9872	Human immune/haema
38	7	4.3	86	22	AAW43671	Human excretory re
39	7	4.3	90	21	AAV16286	Human bladder anti
40	7	4.3	92	22	AAU56894	Eucalyptus grandis
41	7	4.3	103	22	AAW82993	Propionibacterium
42	7	4.3	125	21	AAW828140	Human immune/haema
43	7	4.3	136	22	AAW91688	Sucrose synthase p
44	7	4.3	139	22	ABW08356	Human immune/haema
45	7	4.3	161	22	AAW82089	Novel human diagno

#### ALIGNMENTS

RESULT 1

AAV83951 standard; Protein: 164 AA.

AAV83951:

28-JUL-2000 (first entry)

Human mature interleukin 18 binding protein.

Immunosuppressant; interleukin 18 binding protein; IL18-BP; human; mouse; regulator; drug; sensitivity disease; organ rejection; organ transplant; autoimmune disease.

Homo sapiens.

WO200012555-A1.

09-MAR-2000.

18-NOV-1998; 98WO-JP05186.

01-SEP-1998; 98JP-0247588.

18-NOV-1998; 98JP-0327914.

(HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

Torigoe K, Tanai M, Kurimoto M;

WPT: 2000-237850/20.

N-PSDB: AAA11002.

interleukin 18-binding protein with activity of regulating physiological actions of interleukin 18, useful as regulator and drug

PT for sensitivity diseases and organ rejection and in treating diseases  
 PT due to excess immune reaction -  
 XX Claim 1; Page 44-45; 71pp; Japanese.  
 XX  
 CC The invention relates to novel interleukin 18 (IL-18)-binding proteins  
 CC from humans or mice which act as regulators and drugs for sensitivity  
 CC diseases and organ rejection and in treating diseases due to excess  
 CC immune reaction, e.g. in slowing down rejection after organ transplant,  
 CC and in treating autoimmune diseases. This sequence represents the mature  
 CC human interleukin 12 binding protein.  
 CC  
 XX Sequence 164 AA;  
 SQ  
 Query Match 100.0%; Score 164; DB 21; Length 164;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-158;  
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPVSQTTTAAATASVRSKTPCPSPQPPVFPAAKOCPLAEVTPWEVEVPLNGTSLSCVACS 60  
 DB 1 tpsqgtlaataasvrtskdpcspqppvfpaakcpalevtwpevevplngtislscvacs 60  
 QY 61 RFPNFSILYWLGNFSFIEHLPGRMEGSTSRREGSTGTQCKALVLEQTPALHSTNFSC 120  
 DB 61 rfpnfsilywlgnfsfiehlpgrlwgsgtsrergstgtqckalvleqltpalhstnfsc 120  
 QY 121 VLVDPQVQVORHVLAQLMAGLRATLPPTQEALPSSHSSPOOOG 164  
 DB 121 vlvdpqgvvgrhvlvqlwaglratlptqgealpsshsspqgqg 164  
 RESULT 2  
 AAW98004  
 ID AAW98004 standard; Protein; 192 AA.  
 XX  
 AC AAW98004;  
 XX  
 DT 21-JUN-1999 (first entry)  
 XX  
 DE Human interleukin-18 binding protein splice variant IL-18Bpa.  
 XX  
 KW Interleukin-18 binding protein; IL-18BP; IL-18Bpa; splice variant;  
 KW human; autoimmune disease; inflammation; diabetes; pancreatitis;  
 KW rheumatoid arthritis; graft rejection; sepsis; chronic hepatitis;  
 KW psoriasis; inflammatory bowel disease; multiple sclerosis;  
 KW ischemic heart disease; ischemic brain injury; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 FT /note= "signal peptide"  
 FT Protein 29..192  
 FT /note= "mature protein"  
 XX  
 WO9909063-A1.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 13-AUG-1998; 98WO-IL00379.  
 XX  
 PR 22-JUL-1998; 98IL-0125463.  
 PR 14-AUG-1997; 97IL-0121554.  
 PR 27-AUG-1997; 97IL-0121639.  
 PR 29-SEP-1997; 97IL-0121860.  
 PR 06-NOV-1997; 97IL-0122134.  
 XX  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX  
 PI Dinarello C, Kim SH, Novick D, Rubinstein M;  
 XX  
 DR WPI; 1999-180975/15.

DR N-PSDB; AAX24749.  
 XX  
 PT New interleukin-18 binding protein - useful for treating human  
 PT diseases, including autoimmune disease and inflammation  
 XX  
 PS Claim 3; Page 53-54; 100pp; English.  
 XX  
 CC The present sequence represents a human interleukin-18 binding  
 CC protein (IL-18BP) splice variant designated IL-18Bpa. This  
 CC is one of 4 novel splice variants (see also AAW98005-07), all  
 CC putative soluble proteins, encoded by cDNA clones (see AAX24749-52)  
 CC detected following a screening of human peripheral blood monocyte,  
 CC Jurkat T-cell, peripheral blood mononuclear cell and spleen cell  
 CC cDNA libraries. IL-18Bpa is the most abundant of the 4 splice  
 CC variants and is a putative member of the immunoglobulin superfamily.  
 CC Its first 40 residues perfectly match the N-terminal sequence of  
 CC urinary IL-18BP (see AAW98008). IL-18BP polypeptides capable of  
 CC binding IL-18 and/or modulating and/or blocking IL-18 activity are  
 CC provided. Methods for their isolation and recombinant production,  
 CC DNA vectors expressing them, vectors useful for their expression in  
 CC humans and other mammals, and antibodies against them are also  
 CC provided. IL-18BP polypeptides, and DNA encoding them, can be used  
 CC to treat conditions requiring the protein (claimed). Conditions  
 CC include autoimmune diseases, type I diabetes, rheumatoid arthritis,  
 CC graft rejections, inflammatory bowel disease, sepsis, multiple  
 CC sclerosis, ischemic heart diseases, ischemic brain injury, chronic  
 CC hepatitis, psoriasis, and chronic/acute pancreatitis. IL-18BP is  
 CC also useful for purifying IL-18 (claimed).  
 CC  
 XX Sequence 192 AA;  
 SQ  
 Query Match 100.0%; Score 164; DB 20; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-158;  
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPVSQTTTAAATASVRSKTPCPSPQPPVFPAAKOCPLAEVTPWEVEVPLNGTSLSCVACS 60  
 DB 29 tpsqgtlaataasvrtskdpcspqppvfpaakcpalevtwpevevplngtislscvacs 88  
 QY 61 RFPNFSILYWLGNFSFIEHLPGRMEGSTSRREGSTGTQCKALVLEQTPALHSTNFSC 120  
 DB 89 rfpnfsilywlgnfsfiehlpgrlwgsgtsrergstgtqckalvleqltpalhstnfsc 148  
 QY 121 VLVDPQVQVORHVLAQLMAGLRATLPPTQEALPSSHSSPOOOG 164  
 DB 149 vlvdpqgvvgrhvlvqlwaglratlptqgealpsshsspqgqg 192  
 RESULT 3  
 AAY96292  
 ID AAY96292 standard; protein; 194 AA.  
 XX  
 AC AAY96292;  
 XX  
 DT 16-AUG-2000 (first entry)  
 XX  
 DE Human IGFAM-4 immunoglobulin.  
 XX  
 KW Human; immunoglobulin; IGFAM-4; IGFAM; immune disorder; cancer;  
 KW infection; inflammation; haematopoiesis; AIDS; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..30  
 FT /label= "signal\_peptide"  
 FT Protein 31..194  
 FT /label= "IGFAM-4"  
 XX  
 PN WO200029583-A2.  
 XX  
 PD 25-MAY-2000.

XX 19-NOV-1999; 99WO-US27566.  
XX  
XX  
PR 19-NOV-1998; 99US-0113635.  
PR 22-DEC-1998; 98US-0113635.  
PR 07-APR-1999; 99US-0128194.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Yue H, Tang YT, Corley NC, Guegler KJ, Gorgone GA, Baughn MR;  
PI Lu DAM, Lai P, Hillman JL, Yang J;  
XX  
XX WPI; 2000-387796/33.  
DR N-PSDB; AAA27384.  
XX  
XX Immunoglobulin superfamily proteins, the agonist and antagonist of the  
PT protein is useful for preventing and treating disorders associated with  
PT altered levels of the protein such as cancer, immune system disorders  
XX  
XX  
PS Claim 1; Page 81; 105pp; English.  
XX  
XX The present sequence is the human immunoglobulin superfamily protein  
CC IGFBP-4. Its gene was isolated from a cDNA library of prostate tumour  
CC tissue. It is expressed in nervous, haematopoietic and immune and  
CC cardiovascular tissue, where cancer and inflammation are common. The  
CC gene, protein, its antibodies, agonists and antagonists are suitable for  
CC diagnosing and treating many diseases, including cancer, immune system  
CC disorders (such as inflammation, AIDS, allergies, anaemia,  
CC arteriosclerosis, asthma, atherosclerosis, cholecystitis, Crohn's  
CC disease, diabetes mellitus, emphysema, Graves' disease, hepatitis,  
CC multiple sclerosis, psoriasis, rheumatoid arthritis, scleroderma,  
CC systemic lupus erythematosus and ulcerative colitis), complications of  
CC cancer, haemodialysis and extracorporeal circulation, trauma and  
CC haematopoietic cancer (such as leukaemia) and infections caused by  
CC bacteria, viruses, fungi or parasites.  
XX  
XX  
SQ Sequence 194 AA;

Query Match 100.0%; Score 164; DB 21; Length 194;  
Best Local Similarity 100.0%; Pred. No. 6.2e-158;  
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TPVSOTTTAATAASVSTSDPCSPQPPVFPAAKQCPALVETWPEVPLNGTISTSCVACS 60  
DB 31 TPVSQTLTAATAASVSTKDCPSQPPVFPAAKQCPALVETWPEVPLNGTISTSCVACS 90  
QY 61 RFPNFSILYWLGNCSFIEHLPGRLEWGSTSRERGSTGTOLKALVLEQITPALHSTNFSC 120  
DB 91 RFPNFSILYWLGNCSFIEHLPGRLEWGSTSRERGSTGTOLKALVLEQITPALHSTNFSC 150  
QY 121 VLVDPEQVQVRHVVLAQLWAGLRATLPPTQALPSSHSSPQOQG 164  
DB 151 VLVDPEQVQVRHVVLAQLWAGLRATLPPTQALPSSHSSPQOQG 194

RESULT 4  
AAV83985  
ID AAV83985 standard; Protein; 194 AA.  
XX  
XX AAV83985;  
XX  
XX 28-JUL-2000 (first entry)  
XX  
XX Full length human interleukin 18 binding protein.  
XX  
XX Immunosuppressant; interleukin 18 binding protein; IL18-BP; human; mouse;  
KW regulator; drug; sensitivity disease; organ rejection; organ transplant;  
KW autoimmune disease.  
XX  
XX Homo sapiens.  
XX

FH Key Location/Qualifiers  
FT Peptide 1..30  
FT /label= signal\_peptide  
FT Protein 31..194  
FT /label= "mature human IL-18 binding protein"  
XX  
XX WO200012555-A1.  
XX  
XX 09-MAR-2000.  
XX  
XX 18-NOV-1998; 98WO-JP05186.  
XX  
XX 01-SEP-1998; 98JP-0247588.  
PR 18-NOV-1998; 98JP-0327914.  
XX  
XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
XX  
XX Torioge K, Tanai M, Kurimoto M;  
PI  
XX WPI; 2000-237850/20.  
DR N-PSDB; AAA11007.  
XX  
XX  
XX Example 2; Page 63-64; 71pp; Japanese.  
XX  
XX The invention relates to novel interleukin 18 (IL-18)-binding proteins  
CC from humans or mice which act as regulators and drugs for sensitivity  
CC diseases and organ rejection and in treating diseases due to excess  
CC immune reaction, e.g. in slowing down rejection after organ transplant,  
CC and in treating autoimmune diseases. This sequence represents the full  
CC length human interleukin 12 binding protein.  
XX  
XX  
SQ Sequence 194 AA;

Query Match 100.0%; Score 164; DB 21; Length 194;  
Best Local Similarity 100.0%; Pred. No. 6.2e-158;  
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TPVSOTTTAATAASVSTSDPCSPQPPVFPAAKQCPALVETWPEVPLNGTISTSCVACS 60  
DB 31 TPVSQTLTAATAASVSTKDCPSQPPVFPAAKQCPALVETWPEVPLNGTISTSCVACS 90  
QY 61 RFPNFSILYWLGNCSFIEHLPGRLEWGSTSRERGSTGTOLKALVLEQITPALHSTNFSC 120  
DB 91 RFPNFSILYWLGNCSFIEHLPGRLEWGSTSRERGSTGTOLKALVLEQITPALHSTNFSC 150  
QY 121 VLVDPEQVQVRHVVLAQLWAGLRATLPPTQALPSSHSSPQOQG 164  
DB 151 VLVDPEQVQVRHVVLAQLWAGLRATLPPTQALPSSHSSPQOQG 194

RESULT 5  
AAW98006  
ID AAW98006 standard; Protein; 197 AA.  
XX  
XX AAW98006;  
XX  
XX 21-JUN-1999 (first entry)  
XX  
XX Human interleukin-18 binding protein splice variant IL-18Bpc.  
XX  
XX Interleukin-18 binding protein; IL-18BP; IL-18Bpc; splice variant;  
KW human; autoimmune disease; inflammation; diabetes; pancreatitis;  
KW rheumatoid arthritis; graft rejection; sepsis; chronic hepatitis;  
KW psoriasis; inflammatory bowel disease; multiple sclerosis;  
KW ischemic heart disease; ischemic brain injury; therapy.  
XX  
XX Homo sapiens.  
XX

```

XX Key Location/Qualifiers
FH Peptide 1..28
FT /note= "signal peptide"
FT Protein 29..197
FT /note= "mature protein"
PN WO9909063-A1.
XX
XX 25-FEB-1999.
XX
XX 13-AUG-1998; 98WO-IL00379.
XX
XX 22-JUL-1998; 98IL-0125463.
XX 14-AUG-1997; 97IL-0121554.
XX 27-AUG-1997; 97IL-0121639.
XX 29-SEP-1997; 97IL-0121860.
XX 06-NOV-1997; 97IL-0122134.
XX
XX (YEDA ) YEDA RES & DEV CO LTD.
XX
XX DiNarello C, Kim SH, Novick D, Rubinstein M;
XX WPI; 1999-180975/15.
XX N-PSDB; AAX24751.
XX
XX New interleukin-18 binding protein - useful for treating human
XX diseases, including autoimmune disease and inflammation
XX
XX Claim 3; Page 58-59; 100pp; English.
XX
XX The present sequence represents a human interleukin-18 binding
XX protein (IL-18BP) splice variant designated IL-18Bpc. This is one
XX four novel splice variants (see also AAW98004-07), all putative
XX soluble proteins, that are encoded by cDNA clones (see AAX24749-52)
XX detected following a screening of human peripheral blood monocyte,
XX Jurkat T-cell, peripheral blood mononuclear cell and spleen cell
XX cDNA libraries. IL-18Bpc is a low abundance variant. IL-18BP
XX polypeptides capable of binding IL-18 and/or modulating and/or
XX blocking IL-18 activity are provided, as are methods for their
XX isolation and recombinant production, DNA vectors expressing them,
XX and antibodies against them. IL-18BP polypeptides, and DNA encoding
XX them, can be used to treat conditions requiring the protein
XX (claimed). Conditions include autoimmune diseases, type I diabetes,
XX rheumatoid arthritis, graft rejections, inflammatory bowel disease,
XX sepsis, multiple sclerosis, ischemic heart diseases, ischemic brain
XX injury, chronic hepatitis, psoriasis, and chronic/acute
XX pancreatitis. IL-18BP is also useful for purifying IL-18 (claimed).
XX
XX Sequence 197 AA:
SQ
Query Match 84.8%; Score 139; DB 20; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.3e-132;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPVSQTTTAAATASVRSKPCPSQPPVPPAKKCPALAEVTWPEVEYPLNGTSLSCVACS 60
DB 29 tpsqgttaataasvrskpcpsqppvppakkcpcalaevtwpeveypngtllslscvacs 88
QY 61 RFPNFSILYWLNGSGFIEHLPGRLMGSTSRERGSGTGQLCALVLEQLTPALHSTNFSC 120
DB 89 rfpnfsilywlngsfiehlpgrlmgstsrergsgtgqlckalvleqltpalhstnfsc 148
QY 121 VLVDPEQVQORHVLAQLW 139
DB 149 vlvdpeqvqgrhvnlaqlw 167

```

```

XX AAY83982;
AC 28-JUL-2000 (first entry)
XX
XX Human IL-18 binding protein 5' RACE fragment #1 encoded protein.
DE
XX Immunosuppressant; Interleukin 18 binding protein; IL18-BP; human; mouse;
XX regulator; drug; sensitivity disease; organ rejection; organ transplant;
XX autoimmune disease; rapid amplification of cDNA ends; RACE.
XX
XX Homo sapiens.
XX
XX WO200012555-A1.
XX
XX 09-MAR-2000.
XX
XX 18-NOV-1998; 98WO-JP05186.
XX
XX 01-SEP-1998; 98JP-0247588.
XX 18-NOV-1998; 98JP-0327914.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Toriigoe K, Tanial M, Kurimoto M;
XX WPI; 2000-237850/20.
XX N-PSDB; AAA11004.
XX
XX Interleukin 18-binding protein with activity of regulating
XX physiological actions of Interleukin 18, useful as regulator and drug
XX for sensitivity diseases and organ rejection and in treating diseases
XX due to excess immune reaction -
XX
XX Example 2; Page 60-61; 71pp; Japanese.
XX
XX The invention relates to novel interleukin 18 (IL-18)-binding proteins
XX from humans or mice which act as regulators and drugs for sensitivity
XX diseases and organ rejection and in treating diseases due to excess
XX immune reaction, e.g. in slowing down rejection after organ transplant,
XX and in treating autoimmune diseases. This sequence represents the
XX protein encoded by the initial fragment isolated for a 5' RACE (Rapid
XX amplification of cDNA Ends) reaction for the 5' end of the interleukin
XX 12 binding protein coding sequence.
XX
XX Sequence 137 AA:
SQ
Query Match 83.5%; Score 137; DB 21; Length 137;
Best Local Similarity 100.0%; Pred. No. 1e-130;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPVSQTTTAAATASVRSKPCPSQPPVPPAKKCPALAEVTWPEVEYPLNGTSLSCVACS 60
DB 1 tpsqgttaataasvrskpcpsqppvppakkcpcalaevtwpeveypngtllslscvacs 60
QY 61 RFPNFSILYWLNGSGFIEHLPGRLMGSTSRERGSGTGQLCALVLEQLTPALHSTNFSC 120
DB 61 rfpnfsilywlngsfiehlpgrlmgstsrergsgtgqlckalvleqltpalhstnfsc 120
QY 121 VLVDPEQVQORHVLAQLW 137
DB 121 vlvdpeqvqgrhvnlaqlw 137

```

RESULT 6  
AAW83982  
ID AAY83982 standard; Protein; 137 AA.

RESULT 7  
AAW98007  
ID AAW98007 standard; Protein; 161 AA.  
AC AAW98007;  
XX  
XX 21-JUN-1999 (first entry)  
XX

```

DE Human interleukin-18 binding protein splice variant IL-18BPd.
XX
XX Interleukin-18 binding protein; IL-18BP; IL-18BPd; splice variant;
KM human; autoimmune disease; inflammation; diabetes; pancreatitis;
KM rheumatoid arthritis; graft rejection; sepsis; chronic hepatitis;
KM psoriasis; inflammatory bowel disease; multiple sclerosis;
KM ischemic heart disease; ischemic brain injury; therapy.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..28
FT /note= "signal peptide"
FT 29..161
FT Protein /note= "mature protein"
XX
XX WO9909063-A1.
XX
XX 25-FEB-1999.
XX
XX 13-AUG-1998; 98WO-IL00379.
XX
XX 22-JUL-1998; 98IL-0125463.
XX 14-AUG-1997; 97IL-0121554.
XX 27-AUG-1997; 97IL-0121639.
XX 29-SEP-1997; 97IL-0121860.
XX 06-NOV-1997; 97IL-0122134.
XX
XX (YEDA ) YEDA RES & DEV CO LTD.
XX
XX DiNarello C, Kim SH, Novick D, Rubinstein M;
PI WPI: 1999-180975/15.
XX N-PSDB; AAX24752.
DR
XX New interleukin-18 binding protein - useful for treating human
PT diseases, including autoimmune disease and inflammation
XX
XX Claim 3; Page 60; 100pp; English.
XX
XX The present sequence represents a human interleukin-18 binding
CC protein (IL-18BP) splice variant designated IL-18BPd. This is one
CC four novel splice variants (see also AAW98004-07), all putative
CC soluble proteins, that are encoded by cDNA clones (see AAX24749-52)
CC detected following a screening of human peripheral blood monocyte,
CC Jurkat T-cell, peripheral blood mononuclear cell and spleen cell
CC cDNA libraries. IL-18BPd is a low abundance variant. IL-18BP
CC polypeptides capable of binding IL-18 and/or modulating and/or
CC blocking IL-18 activity are provided, as are methods for their
CC isolation and recombinant production, DNA vectors expressing them,
CC vectors useful for their expression in humans and other mammals,
CC and antibodies against them. IL-18BP polypeptides, and DNA encoding
CC them, can be used to treat conditions requiring the protein
CC (claimed). Conditions include autoimmune diseases, type I diabetes,
CC rheumatoid arthritis, graft rejections, inflammatory bowel disease,
CC sepsis, multiple sclerosis, ischemic heart diseases, ischemic brain
CC injury, chronic hepatitis, psoriasis, and chronic/acute
CC pancreatitis. IL-18BP is also useful for purifying IL-18 (claimed).
XX
XX Sequence 161 AA;
SQ
Query Match 59.1%; Score 97; DB 20; Length 161;
Best Local Similarity 100.0%; Pred. No. 4e-90;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TPVSTTTAATGASVSTDPGSPVPFPAKQCPALFTWTPREVPPLNGTSLSCVACS 60
DB 23 TPVSTTTAATGASVSTDPGSPVPFPAKQCPALFTWTPREVPPLNGTSLSCVACS 88
OY 61 RPPNFSILYWLGNFSTFIEHLPGRLWEGSTSRERGSGTG 97
DB 89 RFPNFSILYWLGNFSTFIEHLPGRLWEGSTSRERGSGTG 125

```

```

RESULT 8
AAW98005
ID AAW98005 standard; Protein, 113 AA.
XX
XX AAW98005;
AC
XX 21-JUN-1999 (first entry)
XX
XX Human interleukin-18 binding protein splice variant IL-18BPd.
DE
XX Interleukin-18 binding protein; IL-18BP; IL-18BPd; splice variant;
KM human; autoimmune disease; inflammation; diabetes; pancreatitis;
KM rheumatoid arthritis; graft rejection; sepsis; chronic hepatitis;
KM psoriasis; inflammatory bowel disease; multiple sclerosis;
KM ischemic heart disease; ischemic brain injury; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..28
FT /note= "signal peptide"
FT 29..113
FT Protein /note= "mature protein"
XX
XX WO9909063-A1.
XX
XX 25-FEB-1999.
XX
XX 13-AUG-1998; 98WO-IL00379.
XX
XX 22-JUL-1998; 98IL-0125463.
XX 14-AUG-1997; 97IL-0121554.
XX 27-AUG-1997; 97IL-0121639.
XX 29-SEP-1997; 97IL-0121860.
XX 06-NOV-1997; 97IL-0122134.
XX
XX (YEDA ) YEDA RES & DEV CO LTD.
XX
XX DiNarello C, Kim SH, Novick D, Rubinstein M;
PI WPI: 1999-180975/15.
XX N-PSDB; AAX24750.
DR
XX New interleukin-18 binding protein - useful for treating human
PT diseases, including autoimmune disease and inflammation
XX
XX Claim 3; Page 55; 100pp; English.
XX
XX The present sequence represents a human interleukin-18 binding
CC protein (IL-18BP) splice variant designated IL-18BPd. This is one
CC four novel splice variants (see also AAW98004-07), all putative
CC soluble proteins, that are encoded by cDNA clones (see AAX24749-52)
CC detected following a screening of human peripheral blood monocyte,
CC Jurkat T-cell, peripheral blood mononuclear cell and spleen cell
CC cDNA libraries. IL-18BPd is a low abundance variant. IL-18BP
CC polypeptides capable of binding IL-18 and/or modulating and/or
CC blocking IL-18 activity are provided, as are methods for their
CC isolation and recombinant production, DNA vectors expressing them,
CC vectors useful for their expression in humans and other mammals,
CC and antibodies against them. IL-18BP polypeptides, and DNA encoding
CC them, can be used to treat conditions requiring the protein
CC (claimed). Conditions include autoimmune diseases, type I diabetes,
CC rheumatoid arthritis, graft rejections, inflammatory bowel disease,
CC sepsis, multiple sclerosis, ischemic heart diseases, ischemic brain
CC injury, chronic hepatitis, psoriasis, and chronic/acute
CC pancreatitis. IL-18BP is also useful for purifying IL-18 (claimed).
XX
XX Sequence 113 AA;
SQ
Query Match 29.3%; Score 48; DB 20; Length 113;

```

Best Local Similarity 100.0%; Pred. No. 1.3e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPVSQTTAATASVRSSTKDCPSQPPVPAKOCPALEVTWPEVEPL 48  
|||||  
Db 29 tpsvqtltaatasvrsstkdcpsqppvfpaakqcpalevtwpevevpl 76

RESULT 9  
AAW83984  
ID AAW83984 standard; Protein; 47 AA.

XX AAW83984;

DT 28-JUL-2000 (first entry)

XX Human IL-18 binding protein C-terminus.

XX Immunosuppressant; Interleukin 18 binding protein; IL18-BP; human; mouse;

XX regulator; drug; sensitivity disease; organ rejection; organ transplant;

XX autoimmune disease; rapid amplification of cDNA ends; RACE.

XX Homo sapiens.

XX WO000012555-A1.

PD 09-MAR-2000.

PF 18-NOV-1998; 98WO-JP05186.

PR 01-SEP-1998; 98JP-0247588.

PR 18-NOV-1998; 98JP-0327914.

XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

XX Torigoe K, Tanai M, Kurimoto M;

XX WPI: 2000-237850/20.

XX N-PSDB: AAA11006.

XX Interleukin 18-binding protein with activity of regulating

XX physiological actions of interleukin 18, useful as regulator and drug

XX for sensitivity diseases and organ rejection and in treating diseases

XX due to excess immune reaction

XX Example 2; Page 62; 71pp; Japanese.

XX The invention relates to novel interleukin 18 (IL-18)-binding proteins

XX from humans or mice which act as regulators and drugs for sensitivity

XX diseases and organ rejection and in treating diseases due to excess

XX immune reaction, e.g. in slowing down rejection after organ transplant,

XX and in treating autoimmune diseases. This sequence represents the

XX C-terminus of the interleukin 18 binding protein encoded by the 3' end

XX of the gene isolated by a 3' RACE (Rapid Amplification of cDNA Ends)

XX reaction.

XX Sequence 47 AA;

XX

XX

XX

XX

XX

XX

XX

XX 21-JUN-1999 (first entry)

XX Human interleukin-18 binding protein N-terminal peptide.

XX Interleukin-18 binding protein; IL-18BP; human; autoimmune disease;

XX inflammation; diabetes; pancreatitis; rheumatoid arthritis;

XX graft rejection; sepsis; chronic hepatitis; psoriasis;

XX inflammatory bowel disease; multiple sclerosis;

XX ischemic heart disease; ischemic brain injury; therapy.

XX Homo sapiens.

XX Key

XX Peptide

XX Protein

XX Location/Qualifiers

XX 1..28

XX /note="signal peptide"

XX 29..113

XX /note="mature protein"

XX WO9909063-A1.

XX 25-FEB-1999.

XX 13-AUG-1998; 98WO-IL00379.

XX 22-JUL-1998; 98IL-0125463.

XX 14-AUG-1997; 97IL-0121554.

XX 27-AUG-1997; 97IL-0121639.

XX 29-SEP-1997; 97IL-0121860.

XX 06-NOV-1997; 97IL-0122134.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX Dinarello C, Kim SH, Novick D, Rubinstein M;

XX WPI: 1999-180975/15.

XX New interleukin-18 binding protein - useful for treating human

XX diseases, including autoimmune disease and inflammation

XX Claim 1; Page 63; 100pp; English.

XX This polypeptide comprises the 40 N-terminal amino acids of human

XX interleukin-18 binding protein (IL-18BP) mature polypeptide. The

XX same 40-amino acid sequence is found at the N-terminus of 4 novel

XX IL-18BP splice variants (see AAW98004-007). IL-18BP polypeptides,

XX including the splice variants, that are capable of binding IL-18

XX and/or modulating and/or blocking IL-18 activity are provided.

XX Methods for their isolation and recombinant production, DNA vectors

XX expressing them, vectors useful for their expression in humans and

XX other mammals, and antibodies against them are also provided.

XX CC IL-18BP polypeptides, and DNA encoding them, can be used to treat

XX conditions requiring the protein (claimed). Such conditions

XX include autoimmune diseases, type I diabetes, rheumatoid arthritis,

XX graft rejections, inflammatory bowel disease, sepsis, multiple

XX sclerosis, ischemic heart diseases, ischemic brain injury, chronic

XX hepatitis, psoriasis, and chronic/acute pancreatitis. IL-18BP is

XX also useful for purifying IL-18 (claimed).

XX Sequence 40 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 24.4%; Score 40; DB 20; Length 40;

Best Local Similarity 100.0%; Pred. No. 6.2e-33;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPVSQTTAATASVRSSTKDCPSQPPVPAKOCPALEVT 40

|||||

Db 1 tpsvqtltaatasvrsstkdcpsqppvfpaakqcpalevt 40

RESULT 11

AAW12716

```
ID  AAY12716 standard; Protein: 55 AA.
XX
AC  AAY12716;
XX
DT  21-JUN-1999 (first entry)
XX
DE  Human 5' EST secreted protein SEQ ID NO:306.
XX
KW  Human: secreted protein; EST; expressed sequence tag; diagnosis;
KW  forensic; gene therapy; chromosome mapping; signal peptide;
KW  upstream regulatory sequence; cytokine activity; cell proliferation;
KW  differentiation; haematopoiesis regulation; tissue growth regulation;
KW  reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW  thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS  Homo sapiens.
XX
PN  WO9906549-A2.
XX
PD  11-FEB-1999.
XX
PF  31-JUL-1998; 98WO-IB01231.
XX
PR  01-AUG-1997; 97US-0905279.
XX
PS  (GEST ) GENSET.
XX
PI  Duclert A, Dumas Milne Edwards J, Lacroix B;
XX  WPI: 1999-153779/13.
XX  N-PSDB: AAX51494.
XX
PT  New nucleic acids encoding human secreted proteins - obtained from
PT  cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
XX
XX  Claim 34: Page 404-405; 522pp; English.
XX
CC  AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
CC  human secreted proteins, and encode the proteins given in AAY1681 to
CC  AAY12913, respectively. The proteins given represent the signal peptide
CC  and an N-terminal fragment of a secreted protein. The nucleic acid
CC  sequences can be used for producing secreted human gene products. They
CC  can also be used to develop products for diagnosis and therapy. The
CC  proteins obtained may have cytokine activity, cell
CC  proliferation/differentiation activity, haematopoiesis regulating
CC  activity, tissue growth regulating activity, reproductive hormone
CC  regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC  thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC  activity, tumour inhibition activity or other activities. The products
CC  can be used in forensic, gene therapy and chromosome mapping procedures.
CC  The sequences can also be used for obtaining corresponding promoter
CC  sequences. The nucleic acids encoding the signal peptide can be used for
CC  directing extracellular secretion of a polypeptide or the insertion of a
CC  polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ  Sequence 55 AA:

Query Match 14.6%; Score 24; DB 20; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.4e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPVSGTTTAAATASVSRSTKDCPPSQ 24
   |||||||||||||||||||||||||
DB 31 lpsqtltaatasvstkdpcpsq 54

RESULT 12
AAY83986
ID AAY83986 standard; Protein: 117 AA.
XX
AC AAY83986;
XX
```

```
DT  28-JUL-2000 (first entry)
XX
XX  Mouse IL-18 binding protein 5' RACE fragment #1 encoded protein.
DE
XX
KW  Immunosuppressant; interleukin 18 binding protein; IL18-BP; human; mouse;
KW  regulator; drug; sensitivity disease; organ rejection; organ transplant;
KW  autoimmune disease; rapid amplification of cDNA ends; RACE.
XX
OS  Mus musculus.
XX
PN  WO200012555-A1.
XX
PD  09-MAR-2000.
XX
PF  18-NOV-1998; 98WO-JP05186.
XX
PR  01-SEP-1998; 98JP-0247588.
XX  18-NOV-1998; 98JP-0327914.
XX
PA  (HAYB ) HAYASHIDARA SEIBUTSU KAKAKU.
XX
PI  Torioge K, Tanai M, Kurimoto M;
XX
DR  WPI: 2000-237850/20.
XX  N-PSDB: AAA11008.
XX
XX  Interleukin 18-binding protein with activity of regulating
PT  physiological actions of interleukin 18, useful as regulator and drug
PT  for sensitivity diseases and organ rejection and in treating diseases
PT  due to excess immune reaction
XX
XX
XX  Example 4: Page 65; 71pp; Japanese.
XX
XX  The invention relates to novel interleukin 18 (IL-18)-binding proteins
XX  from humans or mice which act as regulators and drugs for sensitivity
XX  diseases and organ rejection and in treating diseases due to excess
XX  immune reaction, e.g. in slowing down rejection after organ transplant,
XX  and in treating autoimmune diseases. This sequence represents the
XX  protein encoded by the initial fragment isolated by a 5' RACE (Rapid
XX  Amplification of cDNA Ends) reaction for the 5' end of the mouse
XX  interleukin 12 binding protein coding sequence.
XX
SQ  Sequence 117 AA:

Query Match 12.2%; Score 20; DB 21; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.1e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 FSILYWGNGSFIEHRL 84
   |||||||||||||||||||||
DB 38 fslilywngsfiehlprl 57

RESULT 13
AAY83952
ID AAY83952 standard; Protein: 165 AA.
XX
AC AAY83952;
XX
DT  28-JUL-2000 (first entry)
XX
DE  Mouse mature interleukin 18 binding protein.
XX
XX  Immunosuppressant; interleukin 18 binding protein; IL18-BP; human; mouse;
KW  regulator; drug; sensitivity disease; organ rejection; organ transplant;
KW  autoimmune disease.
XX
OS  Mus musculus.
XX
PN  WO200012555-A1.
XX
PD  09-MAR-2000.
```

```

XX 18-NOV-1998; 98WO-JP05186.
PF 01-SEP-1998; 98JP-0247588.
XX 18-NOV-1998; 98JP-0327914.
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX Torigoe K, Taniai M, Kurimoto M;
XX WPI: 2000-237850/20.
DR N-PSDB; AAA11003.
XX
XX Interleukin 18-binding protein with activity of regulating
XX physiological actions of interleukin 18, useful as regulator and drug
XX for sensitivity diseases and organ rejection and in treating diseases
XX due to excess immune reaction -
XX
XX Claim 1; Page 45; 71pp; Japanese.
XX
XX The invention relates to novel interleukin 18 (IL-18)-binding proteins
XX from humans or mice which act as regulators and drugs for sensitivity
XX diseases and organ rejection and in treating diseases due to excess
XX immune reaction, e.g. in slowing down rejection after organ transplant,
XX and in treating autoimmune diseases. This sequence represents the mature
XX mouse interleukin 12 binding protein.
XX
XX Sequence 165 AA;
SQ

```

```

Query Match 12.2%; Score 20; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 4.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 65 FSILYWLGNCSFIEHLPGRL 84
Db 62 fsilywlgnsgfiehlpgrl 81

```

```

RESULT 14
AA83989
ID AAY83989 standard; Protein; 193 AA.
XX
XX AAY83989;
XX
XX 28-JUL-2000 (first entry)
XX
XX Full length mouse interleukin 18 binding protein.
XX
XX Immunosuppressant; interleukin 18 binding protein; IL18-BP; human; mouse;
XX regulator; drug; sensitivity disease; organ rejection; organ transplant;
XX autoimmune disease.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX Peptide 1..28
XX Protein /label= signal_peptide
XX 29..193
XX /label= "mature IL-18 binding protein"
XX
XX WO200012555-A1.
XX
XX 09-MAR-2000.
XX
XX 18-NOV-1998; 98WO-JP05186.
XX
XX 01-SEP-1998; 98JP-0247588.
XX 18-NOV-1998; 98JP-0327914.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Torigoe K, Taniai M, Kurimoto M;
XX

```

```

XX WPI: 2000-237850/20.
DR N-PSDB; AAA11011.
XX
XX Interleukin 18-binding protein with activity of regulating
XX physiological actions of interleukin 18, useful as regulator and drug
XX for sensitivity diseases and organ rejection and in treating diseases
XX due to excess immune reaction -
XX
XX Example 4; Page 67-69; 71pp; Japanese.
XX
XX The invention relates to novel interleukin 18 (IL-18)-binding proteins
XX from humans or mice which act as regulators and drugs for sensitivity
XX diseases and organ rejection and in treating diseases due to excess
XX immune reaction, e.g. in slowing down rejection after organ transplant,
XX and in treating autoimmune diseases. This sequence represents the full
XX length mouse interleukin 12 binding protein.
XX
XX Sequence 193 AA;
SQ

```

```

Query Match 12.2%; Score 20; DB 21; Length 193;
Best Local Similarity 100.0%; Pred. No. 5e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 65 FSILYWLGNCSFIEHLPGRL 84
Db 90 fsilywlgnsgfiehlpgrl 109

```

```

RESULT 15
AA83983
ID AAY83983 standard; Protein; 49 AA.
XX
XX AAY83983;
XX
XX 28-JUL-2000 (first entry)
XX
XX Human IL-18 binding protein N-terminus.
XX
XX Immunosuppressant; interleukin 18 binding protein; IL18-BP; human; mouse;
XX regulator; drug; sensitivity disease; organ rejection; organ transplant;
XX autoimmune disease; rapid amplification of cDNA ends; RACE.
XX
XX Homo sapiens.
XX
XX WO200012555-A1.
XX
XX 09-MAR-2000.
XX
XX 18-NOV-1998; 98WO-JP05186.
XX
XX 01-SEP-1998; 98JP-0247588.
XX 18-NOV-1998; 98JP-0327914.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Torigoe K, Taniai M, Kurimoto M;
XX
XX WPI: 2000-237850/20.
XX N-PSDB; AAA11005.
XX
XX Interleukin 18-binding protein with activity of regulating
XX physiological actions of interleukin 18, useful as regulator and drug
XX for sensitivity diseases and organ rejection and in treating diseases
XX due to excess immune reaction -
XX
XX Example 2; Page 61-62; 71pp; Japanese.
XX
XX The invention relates to novel interleukin 18 (IL-18)-binding proteins
XX from humans or mice which act as regulators and drugs for sensitivity
XX diseases and organ rejection and in treating diseases due to excess
XX immune reaction, e.g. in slowing down rejection after organ transplant,
XX

```



CC and in treating autoimmune diseases. This sequence represents the  
CC N-terminus of the Interleukin 18 binding protein encoded by the 5' end  
CC of the gene isolated by a 5' RACE (Rapid Amplification of cDNA Ends)  
CC reaction.

XX  
SQ Sequence 49 AA;

Query Match 11.6%; Score 19; DB 21; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1,4e-11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TPVSQTTAATASVRSRKD 19  
|||||  
DB 31 tpvsqtltaatasvrstkd 49

Search completed: October 9, 2002, 01:52:26  
Job time: 5277 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 00:42:39 ; Search time 61.98 Seconds  
(without alignments) 457.747 Million cell updates/sec

Title: US-09-786-130-1

Perfect score: 164  
Sequence: 1 TPVSQTTTAAATASVSRSTKDP.....TLPPQGEALPSSSSPQQQG 164

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPRMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	184	4	Q9UNH2 homo sapien
2	164	100.0	194	4	Q9DBR7 homo sapien
3	97	59.1	161	4	Q9NZAG homo sapien
4	27	16.5	193	11	Q9JUN2 rat
5	20	12.2	191	11	Q9ZOM9 mus musculus
6	20	12.2	192	11	Q9ZOM9 mus musculus
7	20	12.2	193	11	Q9ZOM9 mus musculus
8	19	11.6	218	12	Q9IWI3 molluscum c
9	15	9.1	99	11	Q9CV30 mus musculus
10	11	4.9	421	4	Q9USK4 schizosacch
11	8	4.9	421	4	Q9USK4 schizosacch
12	8	4.9	467	10	Q9A127 zea mays (m
13	8	4.9	505	10	Q9A129 zea mays (m
14	8	4.9	634	5	Q9VDD2 drosophila
15	8	4.9	647	5	Q9VDD2 drosophila
16	8	4.9	647	5	Q9VDD2 drosophila

17	7	4.3	50	2	Q9R4R9	Q9R4R9 bacillus st
18	7	4.3	86	11	Q9JIV3	Q9JIV3 rat
19	7	4.3	167	2	Q935Y1	Q935Y1 synechococ
20	7	4.3	167	16	Q91550	Q91550 pseudomonas
21	7	4.3	207	16	Q9AIX3	Q9AIX3 streptococ
22	7	4.3	240	5	Q9VMA8	Q9VMA8 drosophila
23	7	4.3	246	16	Q9VMA8	Q9VMA8 drosophila
24	7	4.3	263	5	Q9VMA8	Q9VMA8 drosophila
25	7	4.3	310	16	Q9VMA8	Q9VMA8 drosophila
26	7	4.3	330	4	Q14882	Q14882 aquifex aeo
27	7	4.3	339	17	Q9YCC1	Q9YCC1 aeropyrum p
28	7	4.3	356	11	Q9CRS0	Q9CRS0 mus musculu
29	7	4.3	366	5	Q9NKR7	Q9NKR7 drosophila
30	7	4.3	374	15	Q9H694	Q9H694 chimpanzee
31	7	4.3	385	11	Q9ALN9	Q9ALN9 saccharopol
32	7	4.3	401	2	Q9FAR4	Q9FAR4 streptomyce
33	7	4.3	404	2	Q9FAR4	Q9FAR4 streptomyce
34	7	4.3	415	16	Q9A110	Q9A110 streptococ
35	7	4.3	421	13	Q9IBD2	Q9IBD2 brachydanio
36	7	4.3	421	13	Q9YH60	Q9YH60 brachydanio
37	7	4.3	434	2	Q9L409	Q9L409 streptomyce
38	7	4.3	500	4	Q9HAX4	Q9HAX4 homo sapien
39	7	4.3	500	4	Q9HAX3	Q9HAX3 homo sapien
40	7	4.3	500	4	Q9HAX2	Q9HAX2 homo sapien
41	7	4.3	500	16	Q9KBR4	Q9KBR4 bacillus ha
42	7	4.3	502	2	Q9XRO3	Q9XRO3 bacillus st
43	7	4.3	551	10	Q9A169	Q9A169 arabidopsis
44	7	4.3	551	10	Q9A169	Q9A169 arabidopsis
45	7	4.3	554	3	Q9C101	Q9C101 schizosacch

## ALIGNMENTS

RESULT	ID	Query Match	Score	DB	Length	Mismatches	Indels	Gaps
1	Q9UNH2	100.0%	164	4	184	0	0	0
2	Q9UNH2	100.0%	164	4	184	0	0	0
3	Q9UNH2	100.0%	164	4	184	0	0	0
4	Q9UNH2	100.0%	164	4	184	0	0	0
5	Q9UNH2	100.0%	164	4	184	0	0	0
6	Q9UNH2	100.0%	164	4	184	0	0	0
7	Q9UNH2	100.0%	164	4	184	0	0	0
8	Q9UNH2	100.0%	164	4	184	0	0	0
9	Q9UNH2	100.0%	164	4	184	0	0	0
10	Q9UNH2	100.0%	164	4	184	0	0	0
11	Q9UNH2	100.0%	164	4	184	0	0	0
12	Q9UNH2	100.0%	164	4	184	0	0	0
13	Q9UNH2	100.0%	164	4	184	0	0	0
14	Q9UNH2	100.0%	164	4	184	0	0	0
15	Q9UNH2	100.0%	164	4	184	0	0	0
16	Q9UNH2	100.0%	164	4	184	0	0	0
17	Q9UNH2	100.0%	164	4	184	0	0	0
18	Q9UNH2	100.0%	164	4	184	0	0	0
19	Q9UNH2	100.0%	164	4	184	0	0	0
20	Q9UNH2	100.0%	164	4	184	0	0	0
21	Q9UNH2	100.0%	164	4	184	0	0	0
22	Q9UNH2	100.0%	164	4	184	0	0	0
23	Q9UNH2	100.0%	164	4	184	0	0	0
24	Q9UNH2	100.0%	164	4	184	0	0	0
25	Q9UNH2	100.0%	164	4	184	0	0	0
26	Q9UNH2	100.0%	164	4	184	0	0	0
27	Q9UNH2	100.0%	164	4	184	0	0	0
28	Q9UNH2	100.0%	164	4	184	0	0	0
29	Q9UNH2	100.0%	164	4	184	0	0	0
30	Q9UNH2	100.0%	164	4	184	0	0	0
31	Q9UNH2	100.0%	164	4	184	0	0	0
32	Q9UNH2	100.0%	164	4	184	0	0	0
33	Q9UNH2	100.0%	164	4	184	0	0	0
34	Q9UNH2	100.0%	164	4	184	0	0	0
35	Q9UNH2	100.0%	164	4	184	0	0	0
36	Q9UNH2	100.0%	164	4	184	0	0	0
37	Q9UNH2	100.0%	164	4	184	0	0	0
38	Q9UNH2	100.0%	164	4	184	0	0	0
39	Q9UNH2	100.0%	164	4	184	0	0	0
40	Q9UNH2	100.0%	164	4	184	0	0	0
41	Q9UNH2	100.0%	164	4	184	0	0	0
42	Q9UNH2	100.0%	164	4	184	0	0	0
43	Q9UNH2	100.0%	164	4	184	0	0	0
44	Q9UNH2	100.0%	164	4	184	0	0	0
45	Q9UNH2	100.0%	164	4	184	0	0	0

OY 121 VLVDPEQVVRHVLAQLWAGLRATLPPTQELPSSHSSPQOOG 164  
 DB 141 VLVDPEQVVRHVLAQLWAGLRATLPPTQELPSSHSSPQOOG 184

## RESULT 2

Q9UBR7 PRELIMINARY; PRT; 194 AA.  
 AC Q9UBR7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE INTERLEUKIN-18 BINDING PROTEIN PRECURSOR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE=99192308; PubMed=10094485;  
 RA Aizawa Y., Akita K., Tanai M., Korigoe K., Mori T., Nishida Y.,  
 RA Ushio S., Nukada Y., Tanimoto T., Ikegami H., Ikeda M., Kurimoto M.,  
 RT "Cloning and expression of Interleukin-18 binding protein.";  
 RL FEBS Lett. 445:338-342(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99263157; PubMed=10329540;  
 RA Xiang Y., Moss B.,  
 RT "Identification of human and mouse homologs of the M51L-53L-54L  
 RT family of secreted glycoproteins encoded by the Molluscum contagiosum  
 RT poxvirus.";  
 RL Virology 257:297-302(1999).  
 DR EMBL: AB019504; BAA76374.1; -;  
 DR EMBL: AF122906; AADA1051.1; -;  
 DR InterPro: IPR003006; I9\_MHC.  
 DR Pfam: PF00047; I9; 1.  
 KM Signal.  
 FT CHAIN 1 30 POTENTIAL.  
 FT SIGNAL 1 194 INTERLEUKIN-18 BINDING PROTEIN.  
 SO SEQUENCE 194 AA; 21099 MW; 5E520D6E46AFA843 CRC64;

Query Match 100.0%; Score 164; DB 4; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-159;  
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPVSQTTAATASVSTKPCPSQPPVFPAAKQCPALVETWPEVPLNGTSLSCVACS 60  
 DB 31 TPVSQTTAATASVSTKPCPSQPPVFPAAKQCPALVETWPEVPLNGTSLSCVACS 90  
 OY 61 RPPNFSILYWLNGSFIEHLPGRLMEGSTRSGSTGTOLCKALYDEQLTPALHSTNFCSC 120  
 DB 91 RPPNFSILYWLNGSFIEHLPGRLMEGSTRSGSTGTOLCKALYDEQLTPALHSTNFCSC 150  
 OY 121 VLVDPEQVVRHVLAQLWAGLRATLPPTQELPSSHSSPQOOG 164  
 DB 151 VLVDPEQVVRHVLAQLWAGLRATLPPTQELPSSHSSPQOOG 194

RESULT 3  
 Q9NZA9 PRELIMINARY; PRT; 161 AA.  
 AC Q9NZA9;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
 DE INTERLEUKIN-18 BINDING PROTEIN D.  
 GN IL18BP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20122593; PubMed=10655506;  
 RA Kim S.H., Eisenstein M., Reznikov L., Fantuzzi G., Novick D.,  
 RA Rubinstein M., Dinaarello C.A.,  
 RT "Structural requirements of six naturally occurring isoforms of the  
 RT IL-18 binding protein to inhibit IL-18.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1190-1195(2000).  
 DR EMBL: AF215907; AAF31697.1; -;  
 SO SEQUENCE 161 AA; 17246 MW; 28A04CCD92FA833D CRC64;

Query Match 59.1%; Score 97; DB 4; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-91;  
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPVSQTTAATASVSTKPCPSQPPVFPAAKQCPALVETWPEVPLNGTSLSCVACS 60  
 DB 29 TPVSQTTAATASVSTKPCPSQPPVFPAAKQCPALVETWPEVPLNGTSLSCVACS 88  
 OY 61 RPPNFSILYWLNGSFIEHLPGRLMEGSTRSGSTGT 97  
 DB 89 RPPNFSILYWLNGSFIEHLPGRLMEGSTRSGSTGT 125

## RESULT 4

Q9JUN2 PRELIMINARY; PRT; 193 AA.  
 AC Q9JUN2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
 DE INTERLEUKIN 18 BINDING PROTEIN.  
 GN IL18BP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LEWIS;  
 RA Im S.H., Venkatesh N., Barchan D., Souroujon M.C., Fuchs S.,  
 RT "Cloning and characterization of rat IL-18 binding protein.";  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF154569; AAF72102.1; -;  
 SO SEQUENCE 193 AA; 21085 MW; BDB23D80E44B44B0 CRC64;

Query Match 16.5%; Score 27; DB 11; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-19;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 ACSRPPNFSILYWLNGSFIEHLPGRL 84  
 DB 83 ACSRPPNFSILYWLNGSFIEHLPGRL 109

RESULT 5  
 Q9ZOM9 PRELIMINARY; PRT; 191 AA.  
 AC Q9ZOM9;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE INTERLEUKIN-18 BINDING PROTEIN D PRECURSOR.  
 GN IGFBP OR IL18BP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99146382; PubMed=10023777;

```

RA Novick D., Kim S.H., Fantuzzi G., Reznikov L.L., Dinarello C.A.,
RA Rubinstein M.;
RT "Interleukin-18 binding protein: a novel modulator of the Th1 cytokine
RT response.";
RL Immunity 10:127-136(1999).
DR EMBL; AF110803; AAD17194.1; -.
DR MGD; MGI:133800; Iqf1dp.
DR InterPro; IPR003006; Iq_MHC.
DR Pfam; PF00047; Iq; 1.
KW Signal.
FT SIGNAL.
FT CHAIN
FT CHAIN 27 191
SQ SEQUENCE 191 AA; 21035 MW; B2F306A905116ESB CRC64;

Query Match 12.2%; Score 20; DB 11; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 FSILYWLNGSFIEHLPGRL 84
DB 88 FSILYWLNGSFIEHLPGRL 107

RESULT 6
OQ20N0 PRELIMINARY; PRT; 192 AA.
AC OQ20N0;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INTERLEUKIN-18 BINDING PROTEIN C PRECURSOR.
GN IGFBP OR IL18BP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99146382; PubMed=10023777;
RA Novick D., Kim S.H., Fantuzzi G., Reznikov L.L., Dinarello C.A.,
RA Rubinstein M.;
RT "Interleukin-18 binding protein: a novel modulator of the Th1 cytokine
RT response.";
RL Immunity 10:127-136(1999).
DR EMBL; AF110802; AAD17193.1; -.
DR MGD; MGI:133800; Iqf1dp.
DR InterPro; IPR003006; Iq_MHC.
DR Pfam; PF00047; Iq; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 27 192
SQ SEQUENCE 192 AA; 21668 MW; D828CBD328607421 CRC64;

Query Match 12.2%; Score 20; DB 11; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 FSILYWLNGSFIEHLPGRL 84
DB 88 FSILYWLNGSFIEHLPGRL 107

RESULT 7
OQ0UH2 PRELIMINARY; PRT; 193 AA.
AC OQ0UH2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INTERLEUKIN-18 BINDING PROTEIN PRECURSOR (INTERFERON GAMMA INDUCING
DE FACTOR BINDING PROTEIN).

```

```

GN IGFBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99192308; PubMed=10094485;
RA Aizawa Y., Akita K., Tanai M., Korigoe K., Mori T., Nishida Y.,
RA Ushio S., Nukada Y., Tanimoto T., Ikegami H., Ikeda M., Kurimoto M.;
RT "Cloning and expression of interleukin-18 binding protein.";
RL FEBS Lett. 445:338-342(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RX MEDLINE=99263157; PubMed=10329540;
RA Xiang Y., Moss B.;
RT "Identification of human and mouse homologs of the MC5L-53L-54L
RT family of secreted glycoproteins encoded by the Mollusca contagiosum
RT poxvirus.";
RL Virology 257:297-302(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE, EMBRYO, AND SMALL INTESTINE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido T., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Bono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombearts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AB019505; BAA76375.1; -.
DR EMBL; AF122907; AAD41052.1; -.
DR EMBL; AK009721; BAB26462.1; -.
DR EMBL; AK003370; BAB22744.1; -.
DR EMBL; AK008452; BAB25677.1; -.
DR MGD; MGI:133800; Iqf1dp.
DR InterPro; IPR003006; Iq_MHC.
DR Pfam; PF00047; Iq; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 29 193
SQ SEQUENCE 193 AA; 21257 MW; D7AD4B0AB07C8248 CRC64;

Query Match 12.2%; Score 20; DB 11; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 FSILYWLNGSFIEHLPGRL 84
DB 90 FSILYWLNGSFIEHLPGRL 109

RESULT 8
OQ1W13 PRELIMINARY; PRT; 218 AA.
AC OQ1W13;

```

SEQUENCE 290 AA; 32564 MW; 3FB1E295203EAD43 CRC64;

Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 PTOEALPS 155  
Db 235 PTOEALPS 242

RESULT 12  
004127

PRELIMINARY; PRT: 421 AA.

AC 015329;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE SECRETORY MUCIN MUC6 (FRAGMENT).  
GN MUC6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93194895; PubMed=7680650;  
RA Toribara N.W., Robertson A.M., Ho S.B., Kuo W.L., Gum E., Hicks J.W.,  
RA Gum J.R., Byrd J.C., Siddiki B., Kim Y.S.;  
RT "Human gastric mucin. Identification of a unique species by expression  
cloning."  
RL J. Biol. Chem. 268:5879-5885(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97341178; PubMed=9195947;  
RA Toribara N.W., Ho S.B., Gum E., Gum J.R. Jr., Lau P., Kim Y.S.;  
RT "The carboxyl-terminal sequence of the human secretory mucin, MUC6.  
Analysis of the primary amino acid sequence."  
RL J. Biol. Chem. 272:16398-16403(1997).  
DR EMBL: U97698; AAC51370.1; -;  
DR InterPro: IPR000359; Cys\_knot.  
DR Pfam: PF00007; Cys\_knot; 1.  
DR SMART: SM00041; CT\_1.  
DR PROSITE: PS01225; CTCK\_2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 421 AA; 42717 MW; 90E319ED23A9F80 CRC64;

Query Match 4.9%; Score 8; DB 4; Length 421;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 PSSHSPQ 161  
Db 128 PSSHSPQ 135

RESULT 13  
004127

PRELIMINARY; PRT: 467 AA.

AC 004127;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE LETHAL LEAF-SPOT 1 (FRAGMENT).  
GN LLS1.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=PA405 INRED LINE; TISSUE=LEAF;  
RX MEDLINE=97248483; PubMed=9094711;  
RA Gray J., Close P.S., Briggs S.P., Johal G.S.;

RT "A novel suppressor of cell death in plants encoded by the L1s1 gene  
of maize."  
RL Cell 89:25-31(1997).  
DR EMBL: U77345; AAC49676.1; -;  
DR InterPro: IPR001281; Rieske.  
DR Pfam: PF00355; Rieske; 1.  
FT NON\_TER 1  
SQ SEQUENCE 467 AA; 52661 MW; 2D298BA8F4B3B900 CRC64;

Query Match 4.9%; Score 8; DB 10; Length 467;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 PTOEALPS 155  
Db 369 PTOEALPS 376

RESULT 14  
004129

PRELIMINARY; PRT: 505 AA.

AC 004129;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE LETHAL LEAF-SPOT 1 (FRAGMENT).  
GN LLS-1.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=B73 INRED LINE;  
RX MEDLINE=97248483; PubMed=9094711;  
RA Gray J., Close P.S., Briggs S.P., Johal G.S.;  
RT "A novel suppressor of cell death in plants encoded by the L1s1 gene  
of maize."  
RL Cell 89:25-31(1997).  
DR EMBL: U77346; AAC49678.1; -;  
DR InterPro: IPR001281; Rieske.  
DR Pfam: PF00355; Rieske; 1.  
FT NON\_TER 505  
SQ SEQUENCE 505 AA; 56302 MW; A464BA6FEBF0A9 CRC64;

Query Match 4.9%; Score 8; DB 10; Length 505;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 PTOEALPS 155  
Db 422 PTOEALPS 429

RESULT 15  
09VDD2

PRELIMINARY; PRT: 634 AA.

AC 09VDD2;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE SNF4A-GAMMA PROTEIN.  
GN SNF4A-GAMMA OR SNF4AGAMMA OR CG17299.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidae; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003733; AAF55864.1; -;  
 DR FlyBase: FBgn0025803; SNFAA-gamma.  
 DR InterPro: IPR000644; CBS;  
 DR Pfam: PF00571; CBS; 4;  
 DR SMART: SM00116; CBS; 4;  
 SQ SEQUENCE 634 AA; 70174 MM; 3FFD0D53E54BBE7C CRC64;

Query Match 4.9%; Score 8; DB 5; Length 634;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 OTTTAATA 12  
 |||||  
 Db 544 OTTTAATA 551

Search completed: October 9, 2002, 01:53:45  
 Job time: 4266 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 01:47:19 ; Search time 19.28 Seconds  
(without alignments)  
329.357 Million cell updates/sec

Title: US-09-786-130-1  
Perfect score: 164  
Sequence: 1 TPVSOQTTRATATASVSTKDP.....TLPPQGEALPSSHSPQDQG 164

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	84.8	197	118B_HUMAN	Q95998 homo sapien
2	8	4.9	368	1RIBD_SYNY3	Q55158 s riboflavi
3	8	4.9	443	1YAAU_ECOLI	P31679 escherichia
4	8	4.9	838	1AXN2_RAT	Q70240 rattus norv
5	8	4.9	840	1AXN2_MOUSE	Q88566 mus musculu
6	8	4.9	843	1AXN2_HUMAN	Q9Y2E1 homo sapien
7	7	4.3	54	1IOVO_CIRAE	P05579 citricus aeru
8	7	4.3	54	1IOVO_OPHO	P52249 opisthocomu
9	7	4.3	125	1PER_DROAN	Q03293 drosophila
10	7	4.3	353	1HN3G_MOUSE	P35584 mus musculu
11	7	4.3	354	1HN3G_RAT	P32183 rattus norv
12	7	4.3	401	1AMPC_PSYIM	Q05455 psychrobact
13	7	4.3	402	1KAS2_STRCM	P41176 streptomyce
14	7	4.3	500	1PTGI_HUMAN	Q16647 homo sapien
15	7	4.3	552	1EFG1_CANAL	P43064 candida alb
16	7	4.3	600	1DHSA_PARDE	Q59661 paracoccu
17	7	4.3	710	1PAL2_ORYSA	P53443 oryza sativ
18	7	4.3	1017	1HTRA_HUMAN	P54198 homo sapien
19	6	3.7	61	1E306_ADR35	P17531 human adeno
20	6	3.7	65	1MYHB_PIG	P81271 sus scrofa
21	6	3.7	79	1RL11_DICDI	P61618 dictyostell
22	6	3.7	80	1P8_RAT	Q54842 rattus norv
23	6	3.7	100	1KV3C_HUMAN	P01621 homo sapien
24	6	3.7	108	1GLRX_YEAST	P17685 saccharomyc
25	6	3.7	108	1KV3A_HUMAN	P01619 homo sapien
26	6	3.7	109	1KV3B_HUMAN	P01620 homo sapien
27	6	3.7	109	1KV3E_HUMAN	P01622 homo sapien
28	6	3.7	109	1KV3D_HUMAN	P01623 homo sapien
29	6	3.7	109	1KV3G_HUMAN	P04206 homo sapien
30	6	3.7	116	1CUPDA_LOCOMI	P21789 locusta mig
31	6	3.7	118	1DSR4_HUMAN	P56555 homo sapien
32	6	3.7	118	1VE4_HPV13	Q02265 human papil
33	6	3.7	125	1RL11_TOXCA	Q94793 toxocara ca

34	6	3.7	126	1THN1_WHEAT	P01544 triticum ae
35	6	3.7	128	1KV3K_HUMAN	P06311 homo sapien
36	6	3.7	129	1KV3L_HUMAN	P18135 homo sapien
37	6	3.7	129	1KV3M_HUMAN	P18136 homo sapien
38	6	3.7	136	1SR14_HUMAN	P37108 homo sapien
39	6	3.7	136	1H2B1_CHLRE	P50565 chlamydomon
40	6	3.7	153	1Y805_AQUAE	O66989 aquifex aeo
41	6	3.7	155	1UREE_KLEAE	P18317 klebsiella
42	6	3.7	158	1FMA7_BACNO	P27691 bacteroides
43	6	3.7	160	1LACA_FELCA	P21664 felis silve
44	6	3.7	165	1RL11_PIG	Q29205 sus scrofa
45	6	3.7	170	1RL11_CHLRE	P50881 chlamydomon

## ALIGNMENTS

RESULT	ID	STANDARD	PRT	AA
1	118B_HUMAN			
AC	095998: Q96027; Q95993;		197	AA.
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Interleukin-18 binding protein precursor (IL-18BP).			
GN	IL18BP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eulheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99146382; PubMed=10023777;			
RA	Novick D., Kim S.-H., Fantuzzi G., Reznikov L.L., DiNarello C.A.,			
RA	Rubinstein M.;			
RT	"Interleukin-18 binding protein: a novel modulator of the Th1 cytokine			
RT	response.";			
RL	Immunity 10:127-136(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99192308; PubMed=10094485;			
RA	Alzawa Y., Akita K., Tanial M., Torioge K., Mori T., Nishida Y.,			
RA	Ushio S., Nukada Y., Tanimoto T., Ikegami H., Ikeda M., Kurimoto M.;			
RT	"Cloning and expression of interleukin-18 binding protein.";			
RL	FEBS Lett. 445:338-342(1999).			
CC	-1- FUNCTION: FUNCTIONS AS AN INHIBITOR OF THE EARLY TH1 CYTOKINE			
CC	RESPONSE.			
CC	-1- SUBUNIT: BINDS TO IL-18.			
CC	-1- SUBCELLULAR LOCATION: Secreted (Potential).			
CC	-1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A/IL-18BP, B/IL-18BPB AND C/IL-			
CC	18BPB (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: AF110798; AAD17187.1; -			
DR	EMBL: AF110798; AAD17188.1; -			
DR	EMBL: AF110798; AAD17189.1; -			
DR	EMBL: AF110799; AAD17190.1; -			
DR	EMBL: AF110800; AAD17191.1; -			
DR	EMBL: AF110801; AAD17192.1; -			
DR	MIM: 604113; -			
DR	InterPro: IPR003006; Ig_MHC.			
DR	Pfam: PF00047; Ig_1.			
DR	Immunoglobulin domain; Glycoprotein; Signal; Alternative splicing.			
KW	SIGNAL			
FT	1	28	POTENTIAL.	

```

FT CHAIN 29 197 INTERLEUKIN-18 BINDING PROTEIN.
FT DOMAIN 77 155 IG-LIKE C2-TYPE DOMAIN.
FT DISULFID 84 148 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 168 197 VRSPPRGLOEDELCEFHMGKGGLCGSL -> AGLRATL
FT VARSPLIC 77 113 PPTOALPSSSSPPQOG (IN ISOFORM A).
FT VARSPLIC 77 113 NCTLSGVACSRPNSILYGLNGSPFIEHLPGILM
FT VARSPLIC 77 113 SMAEGNLAPHRPSPALDPOOSTAAGLRISTGPAAOP (IN
FT VARSPLIC 77 113 ISOFORM B).
FT VARSPLIC 114 197 MISSING (IN ISOFORM B).
FT VARSPLIC 197 21698 MW; FID626AB2B285E1 CRC64;
SQ SEQUENCE 197 AA; 21698 MW; FID626AB2B285E1 CRC64;

Query Match 84.8%; Score 139; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.4e-135;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPVSGTTTAAATASVSTKPCSPPPVPPAAKCPALLETWEVEVPLNGTSLSCVACS 60
DB 29 TPVSGTTTAAATASVSTKPCSPPPVPPAAKCPALLETWEVEVPLNGTSLSCVACS 88
OY 61 RPNPFSILYWLNGSPFIEHLPGRLMEGSTSRERSTGTQLCAKALYLEQLTPALHSTNFSC 120
DB 89 RPNPFSILYWLNGSPFIEHLPGRLMEGSTSRERSTGTQLCAKALYLEQLTPALHSTNFSC 148
OY 121 VLVDEPQVQVRHVLAQLM 139
DB 149 VLVDEPQVQVRHVLAQLM 167

RESULT 2
RIBD_SYNY3 STANDARD: PRT; 368 AA.
AC 055158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Riboflavin biosynthesis protein ribD [includes:
DE diaminohydroxyphosphoribosylaminoimidine deaminase (EC 3.5.4.26)
DE (Riboflavin-specific deaminase); 5-amino-6-(5-
DE phosphoribosylamino)uracil reductase (EC 1.1.1.193) (HTP reductase)].
GN RIBD OR SLR0066.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneke T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
"Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-diamino-6-hydroxy-4-(5-
CC phosphoribosylamino)pyrimidine + H(2)O = 5-amino-6-(5-
CC phosphoribosylamino)uracil + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-amino-6-(5-phosphoribosylamino)uracil +
CC NADP(+) = 5-amino-6-(5-phosphoribosylamino)uracil + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: D64001; BAA10295.1;
CC InterPro: IPR002734; RibD.C.
CC InterPro: IPR002125; dCMP_cyt_deam.
CC Pfam: PF00383; dCMP_cyt_deam; 1.
CC DR Pfam: PF01872; RibD.C; 1.
CC DR PROSITE: PS00903; CYT_DCM DEAMINASES; 1.
CC KW Riboflavin biosynthesis; Hydroxylase; Zinc; Oxidoreductase; NADP;
CC Multifunctional enzyme; Complete proteome.
FT DOMAIN 1 146 DEAMINASE.
FT DOMAIN 147 368 REDUCTASE.
FT METAL 51 51 ZINC (BY SIMILARITY).
FT METAL 76 76 ZINC (BY SIMILARITY).
FT METAL 85 85 ZINC (BY SIMILARITY).
SQ SEQUENCE 368 AA; 39995 MW; BCD48ED0916B50BC CRC64;

Query Match 4.9%; Score 8; DB 1; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 LVLEQLTP 111
DB 264 LVLEQLTP 271

RESULT 3
YAUU_ECOLI STANDARD: PRT; 443 AA.
AC P31679; P31578; P75628;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical metabolite transport protein yauu.
GN YAUU OR B0045.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
"Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12; MGL655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
"The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC IN POSITION 142 AND 232 THAT PRODUCE TWO SEPARATE ORFS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL: D10483; BAA01321.1; ALT\_FRAME.  
 DR EMBL: AE000114; AAC73156.1; -.  
 DR PIR: S40566; S40566.  
 DR EcoGene: EG11566; yaau.  
 DR InterPro: IPR003662; sub\_transporter.  
 DR Pfam: PF00083; sugar\_tr.1.  
 DR PROSITE: PS00216; SUGAR\_TRANSPORT\_1; FALSE\_NEG.  
 DR PROSITE: PS00217; SUGAR\_TRANSPORT\_2; FALSE\_NEG.  
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;  
 KW Complete proteome.

FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 19 39 1 (POTENTIAL).  
 FT DOMAIN 40 53 PERIPLASMIC (POTENTIAL).  
 FT TRANSSEM 54 74 2 (POTENTIAL).  
 FT DOMAIN 75 84 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 85 105 3 (POTENTIAL).  
 FT DOMAIN 106 113 PERIPLASMIC (POTENTIAL).  
 FT TRANSSEM 114 134 4 (POTENTIAL).  
 FT DOMAIN 135 145 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 146 166 5 (POTENTIAL).  
 FT DOMAIN 167 173 PERIPLASMIC (POTENTIAL).  
 FT TRANSSEM 174 194 6 (POTENTIAL).  
 FT DOMAIN 195 241 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 242 262 7 (POTENTIAL).  
 FT DOMAIN 263 282 PERIPLASMIC (POTENTIAL).  
 FT TRANSSEM 283 303 8 (POTENTIAL).  
 FT DOMAIN 304 309 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 310 329 9 (POTENTIAL).  
 FT DOMAIN 330 334 PERIPLASMIC (POTENTIAL).  
 FT TRANSSEM 335 357 10 (POTENTIAL).  
 FT DOMAIN 358 373 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 374 394 11 (POTENTIAL).  
 FT DOMAIN 395 401 PERIPLASMIC (POTENTIAL).  
 FT TRANSSEM 402 422 12 (POTENTIAL).  
 FT DOMAIN 423 443 CYTOPLASMIC (POTENTIAL).  
 FT TRANSSEM 443 466 13 (POTENTIAL).  
 FT DOMAIN 467 486 14 (POTENTIAL).  
 FT TRANSSEM 487 507 15 (POTENTIAL).  
 FT DOMAIN 508 528 16 (POTENTIAL).  
 FT TRANSSEM 529 549 17 (POTENTIAL).  
 FT DOMAIN 550 570 18 (POTENTIAL).  
 FT TRANSSEM 571 591 19 (POTENTIAL).  
 FT DOMAIN 592 612 20 (POTENTIAL).  
 FT TRANSSEM 613 633 21 (POTENTIAL).  
 FT DOMAIN 634 654 22 (POTENTIAL).  
 FT TRANSSEM 655 675 23 (POTENTIAL).  
 FT DOMAIN 676 696 24 (POTENTIAL).  
 FT TRANSSEM 697 717 25 (POTENTIAL).  
 FT DOMAIN 718 738 26 (POTENTIAL).  
 FT TRANSSEM 739 759 27 (POTENTIAL).  
 FT DOMAIN 760 780 28 (POTENTIAL).  
 FT TRANSSEM 781 801 29 (POTENTIAL).  
 FT DOMAIN 802 822 30 (POTENTIAL).  
 FT TRANSSEM 823 843 31 (POTENTIAL).  
 FT DOMAIN 844 864 32 (POTENTIAL).  
 FT TRANSSEM 865 885 33 (POTENTIAL).  
 FT DOMAIN 886 906 34 (POTENTIAL).  
 FT TRANSSEM 907 927 35 (POTENTIAL).  
 FT DOMAIN 928 948 36 (POTENTIAL).  
 FT TRANSSEM 949 969 37 (POTENTIAL).  
 FT DOMAIN 970 990 38 (POTENTIAL).  
 FT TRANSSEM 991 1011 39 (POTENTIAL).  
 FT DOMAIN 1012 1032 40 (POTENTIAL).  
 FT TRANSSEM 1033 1053 41 (POTENTIAL).  
 FT DOMAIN 1054 1074 42 (POTENTIAL).  
 FT TRANSSEM 1075 1095 43 (POTENTIAL).  
 FT DOMAIN 1096 1116 44 (POTENTIAL).  
 FT TRANSSEM 1117 1137 45 (POTENTIAL).  
 FT DOMAIN 1138 1158 46 (POTENTIAL).  
 FT TRANSSEM 1159 1179 47 (POTENTIAL).  
 FT DOMAIN 1180 1200 48 (POTENTIAL).  
 FT TRANSSEM 1201 1221 49 (POTENTIAL).  
 FT DOMAIN 1222 1242 50 (POTENTIAL).  
 FT TRANSSEM 1243 1263 51 (POTENTIAL).  
 FT DOMAIN 1264 1284 52 (POTENTIAL).  
 FT TRANSSEM 1285 1305 53 (POTENTIAL).  
 FT DOMAIN 1306 1326 54 (POTENTIAL).  
 FT TRANSSEM 1327 1347 55 (POTENTIAL).  
 FT DOMAIN 1348 1368 56 (POTENTIAL).  
 FT TRANSSEM 1369 1389 57 (POTENTIAL).  
 FT DOMAIN 1390 1410 58 (POTENTIAL).  
 FT TRANSSEM 1411 1431 59 (POTENTIAL).  
 FT DOMAIN 1432 1452 60 (POTENTIAL).  
 FT TRANSSEM 1453 1473 61 (POTENTIAL).  
 FT DOMAIN 1474 1494 62 (POTENTIAL).  
 FT TRANSSEM 1495 1515 63 (POTENTIAL).  
 FT DOMAIN 1516 1536 64 (POTENTIAL).  
 FT TRANSSEM 1537 1557 65 (POTENTIAL).  
 FT DOMAIN 1558 1578 66 (POTENTIAL).  
 FT TRANSSEM 1579 1599 67 (POTENTIAL).  
 FT DOMAIN 1600 1620 68 (POTENTIAL).  
 FT TRANSSEM 1621 1641 69 (POTENTIAL).  
 FT DOMAIN 1642 1662 70 (POTENTIAL).  
 FT TRANSSEM 1663 1683 71 (POTENTIAL).  
 FT DOMAIN 1684 1704 72 (POTENTIAL).  
 FT TRANSSEM 1705 1725 73 (POTENTIAL).  
 FT DOMAIN 1726 1746 74 (POTENTIAL).  
 FT TRANSSEM 1747 1767 75 (POTENTIAL).  
 FT DOMAIN 1768 1788 76 (POTENTIAL).  
 FT TRANSSEM 1789 1809 77 (POTENTIAL).  
 FT DOMAIN 1810 1830 78 (POTENTIAL).  
 FT TRANSSEM 1831 1851 79 (POTENTIAL).  
 FT DOMAIN 1852 1872 80 (POTENTIAL).  
 FT TRANSSEM 1873 1893 81 (POTENTIAL).  
 FT DOMAIN 1894 1914 82 (POTENTIAL).  
 FT TRANSSEM 1915 1935 83 (POTENTIAL).  
 FT DOMAIN 1936 1956 84 (POTENTIAL).  
 FT TRANSSEM 1957 1977 85 (POTENTIAL).  
 FT DOMAIN 1978 1998 86 (POTENTIAL).  
 FT TRANSSEM 1999 2019 87 (POTENTIAL).  
 FT DOMAIN 2020 2040 88 (POTENTIAL).  
 FT TRANSSEM 2041 2061 89 (POTENTIAL).  
 FT DOMAIN 2062 2082 90 (POTENTIAL).  
 FT TRANSSEM 2083 2103 91 (POTENTIAL).  
 FT DOMAIN 2104 2124 92 (POTENTIAL).  
 FT TRANSSEM 2125 2145 93 (POTENTIAL).  
 FT DOMAIN 2146 2166 94 (POTENTIAL).  
 FT TRANSSEM 2167 2187 95 (POTENTIAL).  
 FT DOMAIN 2188 2208 96 (POTENTIAL).  
 FT TRANSSEM 2209 2229 97 (POTENTIAL).  
 FT DOMAIN 2230 2250 98 (POTENTIAL).  
 FT TRANSSEM 2251 2271 99 (POTENTIAL).  
 FT DOMAIN 2272 2292 100 (POTENTIAL).  
 FT TRANSSEM 2293 2313 101 (POTENTIAL).  
 FT DOMAIN 2314 2334 102 (POTENTIAL).  
 FT TRANSSEM 2335 2355 103 (POTENTIAL).  
 FT DOMAIN 2356 2376 104 (POTENTIAL).  
 FT TRANSSEM 2377 2397 105 (POTENTIAL).  
 FT DOMAIN 2398 2418 106 (POTENTIAL).  
 FT TRANSSEM 2419 2439 107 (POTENTIAL).  
 FT DOMAIN 2440 2460 108 (POTENTIAL).  
 FT TRANSSEM 2461 2481 109 (POTENTIAL).  
 FT DOMAIN 2482 2502 110 (POTENTIAL).  
 FT TRANSSEM 2503 2523 111 (POTENTIAL).  
 FT DOMAIN 2524 2544 112 (POTENTIAL).  
 FT TRANSSEM 2545 2565 113 (POTENTIAL).  
 FT DOMAIN 2566 2586 114 (POTENTIAL).  
 FT TRANSSEM 2587 2607 115 (POTENTIAL).  
 FT DOMAIN 2608 2628 116 (POTENTIAL).  
 FT TRANSSEM 2629 2649 117 (POTENTIAL).  
 FT DOMAIN 2650 2670 118 (POTENTIAL).  
 FT TRANSSEM 2671 2691 119 (POTENTIAL).  
 FT DOMAIN 2692 2712 120 (POTENTIAL).  
 FT TRANSSEM 2713 2733 121 (POTENTIAL).  
 FT DOMAIN 2734 2754 122 (POTENTIAL).  
 FT TRANSSEM 2755 2775 123 (POTENTIAL).  
 FT DOMAIN 2776 2796 124 (POTENTIAL).  
 FT TRANSSEM 2797 2817 125 (POTENTIAL).  
 FT DOMAIN 2818 2838 126 (POTENTIAL).  
 FT TRANSSEM 2839 2859 127 (POTENTIAL).  
 FT DOMAIN 2860 2880 128 (POTENTIAL).  
 FT TRANSSEM 2881 2901 129 (POTENTIAL).  
 FT DOMAIN 2902 2922 130 (POTENTIAL).  
 FT TRANSSEM 2923 2943 131 (POTENTIAL).  
 FT DOMAIN 2944 2964 132 (POTENTIAL).  
 FT TRANSSEM 2965 2985 133 (POTENTIAL).  
 FT DOMAIN 2986 3006 134 (POTENTIAL).  
 FT TRANSSEM 3007 3027 135 (POTENTIAL).  
 FT DOMAIN 3028 3048 136 (POTENTIAL).  
 FT TRANSSEM 3049 3069 137 (POTENTIAL).  
 FT DOMAIN 3070 3090 138 (POTENTIAL).  
 FT TRANSSEM 3091 3111 139 (POTENTIAL).  
 FT DOMAIN 3112 3132 140 (POTENTIAL).  
 FT TRANSSEM 3133 3153 141 (POTENTIAL).  
 FT DOMAIN 3154 3174 142 (POTENTIAL).  
 FT TRANSSEM 3175 3195 143 (POTENTIAL).  
 FT DOMAIN 3196 3216 144 (POTENTIAL).  
 FT TRANSSEM 3217 3237 145 (POTENTIAL).  
 FT DOMAIN 3238 3258 146 (POTENTIAL).  
 FT TRANSSEM 3259 3279 147 (POTENTIAL).  
 FT DOMAIN 3280 3300 148 (POTENTIAL).  
 FT TRANSSEM 3301 3321 149 (POTENTIAL).  
 FT DOMAIN 3322 3342 150 (POTENTIAL).  
 FT TRANSSEM 3343 3363 151 (POTENTIAL).  
 FT DOMAIN 3364 3384 152 (POTENTIAL).  
 FT TRANSSEM 3385 3405 153 (POTENTIAL).  
 FT DOMAIN 3406 3426 154 (POTENTIAL).  
 FT TRANSSEM 3427 3447 155 (POTENTIAL).  
 FT DOMAIN 3448 3468 156 (POTENTIAL).  
 FT TRANSSEM 3469 3489 157 (POTENTIAL).  
 FT DOMAIN 3490 3510 158 (POTENTIAL).  
 FT TRANSSEM 3511 3531 159 (POTENTIAL).  
 FT DOMAIN 3532 3552 160 (POTENTIAL).  
 FT TRANSSEM 3553 3573 161 (POTENTIAL).  
 FT DOMAIN 3574 3594 162 (POTENTIAL).  
 FT TRANSSEM 3595 3615 163 (POTENTIAL).  
 FT DOMAIN 3616 3636 164 (POTENTIAL).  
 FT TRANSSEM 3637 3657 165 (POTENTIAL).  
 FT DOMAIN 3658 3678 166 (POTENTIAL).  
 FT TRANSSEM 3679 3699 167 (POTENTIAL).  
 FT DOMAIN 3700 3720 168 (POTENTIAL).  
 FT TRANSSEM 3721 3741 169 (POTENTIAL).  
 FT DOMAIN 3742 3762 170 (POTENTIAL).  
 FT TRANSSEM 3763 3783 171 (POTENTIAL).  
 FT DOMAIN 3784 3804 172 (POTENTIAL).  
 FT TRANSSEM 3805 3825 173 (POTENTIAL).  
 FT DOMAIN 3826 3846 174 (POTENTIAL).  
 FT TRANSSEM 3847 3867 175 (POTENTIAL).  
 FT DOMAIN 3868 3888 176 (POTENTIAL).  
 FT TRANSSEM 3889 3909 177 (POTENTIAL).  
 FT DOMAIN 3910 3930 178 (POTENTIAL).  
 FT TRANSSEM 3931 3951 179 (POTENTIAL).  
 FT DOMAIN 3952 3972 180 (POTENTIAL).  
 FT TRANSSEM 3973 3993 181 (POTENTIAL).  
 FT DOMAIN 3994 4014 182 (POTENTIAL).  
 FT TRANSSEM 4015 4035 183 (POTENTIAL).  
 FT DOMAIN 4036 4056 184 (POTENTIAL).  
 FT TRANSSEM 4057 4077 185 (POTENTIAL).  
 FT DOMAIN 4078 4098 186 (POTENTIAL).  
 FT TRANSSEM 4099 4119 187 (POTENTIAL).  
 FT DOMAIN 4120 4140 188 (POTENTIAL).  
 FT TRANSSEM 4141 4161 189 (POTENTIAL).  
 FT DOMAIN 4162 4182 190 (POTENTIAL).  
 FT TRANSSEM 4183 4203 191 (POTENTIAL).  
 FT DOMAIN 4204 4224 192 (POTENTIAL).  
 FT TRANSSEM 4225 4245 193 (POTENTIAL).  
 FT DOMAIN 4246 4266 194 (POTENTIAL).  
 FT TRANSSEM 4267 4287 195 (POTENTIAL).  
 FT DOMAIN 4288 4308 196 (POTENTIAL).  
 FT TRANSSEM 4309 4329 197 (POTENTIAL).  
 FT DOMAIN 4330 4350 198 (POTENTIAL).  
 FT TRANSSEM 4351 4371 199 (POTENTIAL).  
 FT DOMAIN 4372 4392 200 (POTENTIAL).  
 FT TRANSSEM 4393 4413 201 (POTENTIAL).  
 FT DOMAIN 4414 4434 202 (POTENTIAL).  
 FT TRANSSEM 4435 4455 203 (POTENTIAL).  
 FT DOMAIN 4456 4476 204 (POTENTIAL).  
 FT TRANSSEM 4477 4497 205 (POTENTIAL).  
 FT DOMAIN 4498 4518 206 (POTENTIAL).  
 FT TRANSSEM 4519 4539 207 (POTENTIAL).  
 FT DOMAIN 4540 4560 208 (POTENTIAL).  
 FT TRANSSEM 4561 4581 209 (POTENTIAL).  
 FT DOMAIN 4582 4602 210 (POTENTIAL).  
 FT TRANSSEM 4603 4623 211 (POTENTIAL).  
 FT DOMAIN 4624 4644 212 (POTENTIAL).  
 FT TRANSSEM 4645 4665 213 (POTENTIAL).  
 FT DOMAIN 4666 4686 214 (POTENTIAL).  
 FT TRANSSEM 4687 4707 215 (POTENTIAL).  
 FT DOMAIN 4708 4728 216 (POTENTIAL).  
 FT TRANSSEM 4729 4749 217 (POTENTIAL).  
 FT DOMAIN 4750 4770 218 (POTENTIAL).  
 FT TRANSSEM 4771 4791 219 (POTENTIAL).  
 FT DOMAIN 4792 4812 220 (POTENTIAL).  
 FT TRANSSEM 4813 4833 221 (POTENTIAL).  
 FT DOMAIN 4834 4854 222 (POTENTIAL).  
 FT TRANSSEM 4855 4875 223 (POTENTIAL).  
 FT DOMAIN 4876 4896 224 (POTENTIAL).  
 FT TRANSSEM 4897 4917 225 (POTENTIAL).  
 FT DOMAIN 4918 4938 226 (POTENTIAL).  
 FT TRANSSEM 4939 4959 227 (POTENTIAL).  
 FT DOMAIN 4960 4980 228 (POTENTIAL).  
 FT TRANSSEM 4981 5001 229 (POTENTIAL).  
 FT DOMAIN 5002 5022 230 (POTENTIAL).  
 FT TRANSSEM 5023 5043 231 (POTENTIAL).  
 FT DOMAIN 5044 5064 232 (POTENTIAL).  
 FT TRANSSEM 5065 5085 233 (POTENTIAL).  
 FT DOMAIN 5086 5106 234 (POTENTIAL).  
 FT TRANSSEM 5107 5127 235 (POTENTIAL).  
 FT DOMAIN 5128 5148 236 (POTENTIAL).  
 FT TRANSSEM 5149 5169 237 (POTENTIAL).  
 FT DOMAIN 5170 5190 238 (POTENTIAL).  
 FT TRANSSEM 5191 5211 239 (POTENTIAL).  
 FT DOMAIN 5212 5232 240 (POTENTIAL).  
 FT TRANSSEM 5233 5253 241 (POTENTIAL).  
 FT DOMAIN 5254 5274 242 (POTENTIAL).  
 FT TRANSSEM 5275 5295 243 (POTENTIAL).  
 FT DOMAIN 5296 5316 244 (POTENTIAL).  
 FT TRANSSEM 5317 5337 245 (POTENTIAL).  
 FT DOMAIN 5338 5358 246 (POTENTIAL).  
 FT TRANSSEM 5359 5379 247 (POTENTIAL).  
 FT DOMAIN 5380 5400 248 (POTENTIAL).  
 FT TRANSSEM 5401 5421 249 (POTENTIAL).  
 FT DOMAIN 5422 5442 250 (POTENTIAL).  
 FT TRANSSEM 5443 5463 251 (POTENTIAL).  
 FT DOMAIN 5464 5484 252 (POTENTIAL).  
 FT TRANSSEM 5485 5505 253 (POTENTIAL).  
 FT DOMAIN 5506 5526 254 (POTENTIAL).  
 FT TRANSSEM 5527 5547 255 (POTENTIAL).  
 FT DOMAIN 5548 5568 256 (POTENTIAL).  
 FT TRANSSEM 5569 5589 257 (POTENTIAL).  
 FT DOMAIN 5590 5610 258 (POTENTIAL).  
 FT TRANSSEM 5611 5631 259 (POTENTIAL).  
 FT DOMAIN 5632 5652 260 (POTENTIAL).  
 FT TRANSSEM 5653 5673 261 (POTENTIAL).  
 FT DOMAIN 5674 5694 262 (POTENTIAL).  
 FT TRANSSEM 5695 5715 263 (POTENTIAL).  
 FT DOMAIN 5716 5736 264 (POTENTIAL).  
 FT TRANSSEM 5737 5757 265 (POTENTIAL).  
 FT DOMAIN 5758 5778 266 (POTENTIAL).  
 FT TRANSSEM 5779 5799 267 (POTENTIAL).  
 FT DOMAIN 5800 5820 268 (POTENTIAL).  
 FT TRANSSEM 5821 5841 269 (POTENTIAL).  
 FT DOMAIN 5842 5862 270 (POTENTIAL).  
 FT TRANSSEM 5863 5883 271 (POTENTIAL).  
 FT DOMAIN 5884 5904 272 (POTENTIAL).  
 FT TRANSSEM 5905 5925 273 (POTENTIAL).  
 FT DOMAIN 5926 5946 274 (POTENTIAL).  
 FT TRANSSEM 5947 5967 275 (POTENTIAL).  
 FT DOMAIN 5968 5988 276 (POTENTIAL).  
 FT TRANSSEM 5989 6009 277 (POTENTIAL).  
 FT DOMAIN 6010 6030 278 (POTENTIAL).  
 FT TRANSSEM 6031 6051 279 (POTENTIAL).  
 FT DOMAIN 6052 6072 280 (POTENTIAL).  
 FT TRANSSEM 6073 6093 281 (POTENTIAL).  
 FT DOMAIN 6094 6114 282 (POTENTIAL).  
 FT TRANSSEM 6115 6135 283 (POTENTIAL).  
 FT DOMAIN 6136 6156 284 (POTENTIAL).  
 FT TRANSSEM 6157 6177 285 (POTENTIAL).  
 FT DOMAIN 6178 6198 286 (POTENTIAL).  
 FT TRANSSEM 6199 6219 287 (POTENTIAL).  
 FT DOMAIN 6220 6240 288 (POTENTIAL).  
 FT TRANSSEM 6241 6261 289 (POTENTIAL).  
 FT DOMAIN 6262 6282 290 (POTENTIAL).  
 FT TRANSSEM 6283 6303 291 (POTENTIAL).  
 FT DOMAIN 6304 6324 292 (POTENTIAL).  
 FT TRANSSEM 6325 6345 293 (POTENTIAL).  
 FT DOMAIN 6346 6366 294 (POTENTIAL).  
 FT TRANSSEM 6367 6387 295 (POTENTIAL).  
 FT DOMAIN 6388 6408 296 (POTENTIAL).  
 FT TRANSSEM 6409 6429 297 (POTENTIAL).  
 FT DOMAIN 6430 6450 298 (POTENTIAL).  
 FT TRANSSEM 6451 6471 299 (POTENTIAL).  
 FT DOMAIN 6472 6492 300 (POTENTIAL).  
 FT TRANSSEM 6493 6513 301 (POTENTIAL).  
 FT DOMAIN 6514 6534 302 (POTENTIAL).  
 FT TRANSSEM 6535 6555 303 (POTENTIAL).  
 FT DOMAIN 6556 6576 304 (POTENTIAL).  
 FT TRANSSEM 6577 6597 305 (POTENTIAL).  
 FT DOMAIN 6598 6618 306 (POTENTIAL).  
 FT TRANSSEM 6619 6639 307 (POTENTIAL).  
 FT DOMAIN 6640 6660 308 (POTENTIAL).  
 FT TRANSSEM 6661 6681 309 (POTENTIAL).  
 FT DOMAIN 6682 6702 310 (POTENTIAL).  
 FT TRANSSEM 6703 6723 311 (POTENTIAL).  
 FT DOMAIN 6724 6744 312 (POTENTIAL).  
 FT TRANSSEM 6745 6765 313 (POTENTIAL).  
 FT DOMAIN 6766 6786 314 (POTENTIAL).  
 FT TRANSSEM 6787 6807 315 (POTENTIAL).  
 FT DOMAIN 6808 6828 316 (POTENTIAL).  
 FT TRANSSEM 6829 6849 317 (POTENTIAL).  
 FT DOMAIN 6850 6870 318 (POTENTIAL).  
 FT TRANSSEM 6871 6891 319 (POTENTIAL).  
 FT DOMAIN 6892 6912 320 (POTENTIAL).  
 FT TRANSSEM 6913 6933 321 (POTENTIAL).  
 FT DOMAIN 6934 6954 322 (POTENTIAL).  
 FT TRANSSEM 6955 6975 323 (POTENTIAL).  
 FT DOMAIN 6976 6996 324 (POTENTIAL).  
 FT TRANSSEM 6997 7017 325 (POTENTIAL).  
 FT DOMAIN 7018 7038 326 (POTENTIAL).  
 FT TRANSSEM 7039 7059 327 (POTENTIAL).  
 FT DOMAIN 7060 7080 328 (POTENTIAL).  
 FT TRANSSEM 7081 7101 329 (POTENTIAL).  
 FT DOMAIN 7102 7122 330 (POTENTIAL).  
 FT TRANSSEM 7123 7143 331 (POTENTIAL).  
 FT DOMAIN 7144 7164 332 (POTENTIAL).  
 FT TRANSSEM 7165 7185 333 (POTENTIAL).  
 FT DOMAIN 7186 7206 334 (POTENTIAL).  
 FT TRANSSEM 7207 7227 335 (POTENTIAL).  
 FT DOMAIN 7228 7248 336 (POTENTIAL).  
 FT TRANSSEM 7249 7269 337 (POTENTIAL).  
 FT DOMAIN 7270 7290 338 (POTENTIAL).  
 FT TRANSSEM 7291 7311 339 (POTENTIAL).  
 FT DOMAIN 7312 7332 340 (POTENTIAL).  
 FT TRANSSEM 7333 7353 341 (POTENTIAL).  
 FT DOMAIN 7354 7374 342 (POTENTIAL).  
 FT TRANSSEM 7375 7395 343 (POTENTIAL).  
 FT DOMAIN 7396 7416 344 (POTENTIAL).  
 FT TRANSSEM 7417 7437 345 (POTENTIAL).  
 FT DOMAIN 7438 7458 346 (POTENTIAL).  
 FT TRANSSEM 7459 7479 347 (POTENTIAL).  
 FT DOMAIN 7480 7500 348 (POTENTIAL).  
 FT TRANSSEM 7501 7521 349 (POTENTIAL).  
 FT DOMAIN 7522 7542 350 (POTENTIAL).  
 FT TRANSSEM 7543 7563 351 (POTENTIAL).  
 FT DOMAIN 7564 7584 352 (POTENTIAL).  
 FT TRANSSEM 7585 7605 353 (POTENTIAL).  
 FT DOMAIN 7606 7626 354 (POTENTIAL).  
 FT TRANSSEM 7627 7647 355 (POTENTIAL).  
 FT DOMAIN 7648 7668 356 (POTENTIAL).  
 FT TRANSSEM 7669 7689 357 (POTENTIAL).  
 FT DOMAIN 7690 7710 358 (POTENTIAL).  
 FT TRANSSEM 7711 7731 359 (POTENTIAL).  
 FT DOMAIN 7732 7752 360 (POTENTIAL).  
 FT TRANSSEM 7753 7773 361 (POTENTIAL).  
 FT DOMAIN 7774 7794 362 (POTENTIAL).  
 FT TRANSSEM 7795 7815 363 (POTENTIAL).  
 FT DOMAIN 7816 7836 364 (POTENTIAL).  
 FT TRANSSEM 7837 7857 365 (POTENTIAL).  
 FT DOMAIN 7858 7878 366 (POTENTIAL).  
 FT TRANSSEM 7879 7899 367 (POTENTIAL).  
 FT DOMAIN 7900 7920 368 (POTENTIAL).  
 FT TRANSSEM 7921 7941 369 (POTENTIAL).  
 FT DOMAIN 7942 7962 370 (POTENTIAL).  
 FT TRANSSEM 7963 7983 371 (POTENTIAL).  
 FT DOMAIN 7984 8004 372 (POTENTIAL).  
 FT TRANSSEM 8005 8025 373 (POTENTIAL).  
 FT DOMAIN 8026 8046 374 (POTENTIAL).  
 FT TRANSSEM 8047 8067 375 (POTENTIAL).  
 FT DOMAIN 8068 8088 376 (POTENTIAL).  
 FT TRANSSEM 8089 8109 377 (POTENTIAL).  
 FT DOMAIN 8110 8130 378 (POTENTIAL).  
 FT TRANSSEM 8131 8151 379 (POTENTIAL).  
 FT DOMAIN 8152 8172 380 (POTENTIAL).  
 FT TRANSSEM 8173 8193 381 (POTENTIAL).  
 FT DOMAIN 8194 8214 382 (POTENTIAL).  
 FT TRANSSEM 8215 8235 383 (POTENTIAL).  
 FT DOMAIN 8236 8256 384 (POTENTIAL).  
 FT TRANSSEM 8257 8277 385 (POTENTIAL).  
 FT DOMAIN 8278 8298 386 (POTENTIAL).  
 FT TRANSSEM 8299 8319 387 (POTENTIAL).  
 FT DOMAIN 8320 8340 388 (POTENTIAL).  
 FT TRANSSEM 8341 8361 389 (POTENTIAL).  
 FT DOMAIN 8362 8382 390 (POTENTIAL).  
 FT TRANSSEM 8383 8403 391 (POTENTIAL).  
 FT DOMAIN 8404 8424 392 (POTENTIAL).  
 FT TRANSSEM 8425 8445 393 (POTENTIAL).  
 FT DOMAIN 8446 8466 394 (POTENTIAL).  
 FT TRANSSEM 8467 8487 395 (POTENTIAL).  
 FT DOMAIN 8488 8508 396 (POTENTIAL).  
 FT TRANSSEM 8509 8529 397 (POTENTIAL).  
 FT DOMAIN 8530 8550 398 (POTENTIAL).  
 FT TRANSSEM 8551 8571 399 (POTENTIAL).  
 FT DOMAIN 8572 8592 400 (POTENTIAL).  
 FT TRANSSEM 8593 8613 401 (POTENTIAL).  
 FT DOMAIN 8614 8634 402 (POTENTIAL).  
 FT TRANSSEM 8635 8655 403 (POTENTIAL).  
 FT DOMAIN 8656 8676 404 (POTENTIAL).  
 FT TRANSSEM 8677 8697 405 (POTENTIAL).  
 FT DOMAIN 8698 8718 406 (POTENTIAL).  
 FT TRANSSEM 8719 8739 407 (POTENTIAL).  
 FT DOMAIN 8740 8760 408 (POTENTIAL).  
 FT TRANSSEM 8761 8781 409 (POTENTIAL).  
 FT DOMAIN 8782 8802 410 (POTENTIAL).  
 FT TRANSSEM 8803 8823 411 (POTENTIAL).  
 FT DOMAIN 8824 8844 412 (POTENTIAL).  
 FT TRANSSEM 8845 8865 413 (POTENTIAL).  
 FT DOMAIN 8866 8886 414 (POTENTIAL).  
 FT TRANSSEM 8887 8907 415 (POTENTIAL).  
 FT DOMAIN 8908 8928 416 (POTENTIAL).  
 FT TRANSSEM 8929 8949 417 (POTENTIAL).  
 FT DOMAIN 8950 8970 418 (POTENTIAL).  
 FT TRANSSEM 8971 8991 419 (POTENTIAL).  
 FT DOMAIN 8992 9012 420 (POTENTIAL).  
 FT TRANSSEM 9013 9033 421 (POTENTIAL).  
 FT DOMAIN 9034 9054 422 (POTENTIAL).  
 FT TRANSSEM 9055 9075 423 (POTENTIAL).  
 FT DOMAIN 9076 9096 424 (POTENTIAL).  
 FT TRANSSEM 9097 9117 425 (POTENTIAL).  
 FT DOMAIN 9118 9138 426 (POTENTIAL).  
 FT TRANSSEM 9139 9159 427 (POTENTIAL).  
 FT DOMAIN 9160 9180 428 (POTENTIAL).  
 FT TRANSSEM 9181 9201 429 (POTENTIAL).  
 FT DOMAIN 9202 9222 430 (POTENTIAL).  
 FT TRANSSEM 9223 9243 431 (POTENTIAL).  
 FT DOMAIN 9244 9264 432 (POTENTIAL).  
 FT TRANSSEM 9265 9285 433 (POTENTIAL).  
 FT DOMAIN 9286 9306 434 (POTENTIAL).  
 FT TRANSSEM 9307 9327 435 (POTENTIAL).  
 FT DOMAIN 9328 9348 436 (POTENTIAL).  
 FT TRANSSEM 9349 9369 437 (POTENTIAL).  
 FT DOMAIN 9370 9390 438 (POTENTIAL).  
 FT TRANSSEM 9391 9411 439 (POTENTIAL).  
 FT DOMAIN 9412 9432 440 (POTENTIAL).  
 FT TRANSSEM 9433 9453 441 (POTENTIAL).  
 FT DOMAIN 9454 9474 442 (POTENTIAL).  
 FT TRANSSEM 9475 9495 443 (POTENTIAL).  
 FT DOMAIN 9496 9

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-98221239; PubMed-9554852;  
 RA Behrens J., Jerchow B.-A., Muertels M., Grimm J., Asbrand C.,  
 RA Wirtz R., Kuehl M., Wedlich D., Birchmeier W.;  
 RT "Functional interaction of an axin homolog, conductin, with beta-  
 RT catenin, APC, and GSK3beta.";  
 RL Science 280:596-599(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Zhang T., Fagotto F., Hsu W., Zeng L., Gilbert D., Copeland N.G.,  
 RA Jenkins N.A., Warburton D., Costantini F.;  
 RT "Properties of mouse Axin2 and human AXIN2: chromosomal location,  
 RT expression pattern, interaction with Axin and effects on embryonic  
 RT axis formation.";  
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
 CC -I- FUNCTION: INHIBITOR OF THE WNT SIGNALING PATHWAY. DOWN-REGULATES  
 CC BETA-CATENIN. PROBABLY FACILITATE THE PHOSPHORYLATION OF BETA-  
 CC CATEININ AND APC BY GSK-3B (BY SIMILARITY).  
 CC -I- SUBUNIT: INTERACTS WITH GLYCOGEN SYNTHASE KINASE-3 BETA (GSK-3B)  
 CC AND BETA-CATENIN. THE INTERACTION BETWEEN AXIN AND BETA-CATENIN  
 CC OCCURS VIA THE ARMADILLO REPEATS CONTAINED IN BETA-CATENIN.  
 CC TERNARY COMPLEX (BY SIMILARITY).  
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -I- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B AND DEPHOSPHORYLATED BY  
 CC PP2A (BY SIMILARITY).  
 CC -I- SIMILARITY: CONTAINS 1 RGS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF0733788; AAC26047.1; -;  
 DR EMBL: AF205889; AAF22800.1; -;  
 DR HSSP: P49799; IAGR.  
 DR MGD: MGI:1270862; Axin2.  
 DR InterPro: IPR001158; DIX.  
 DR InterPro: IPR000342; RGS.  
 DR Pfam: PF00778; DIX; 1.  
 DR Pfam: PF00615; RGS; 1.  
 DR PRINTS: PR01301; RGS-PROTEIN.  
 DR PRODOM: PD001580; RGS; 1.  
 DR Prodom: PD003639; DIX; 1.  
 DR SMART: SM00021; DAX; 1.  
 DR SMART: SM00315; RGS; 1.  
 DR PROSITE: PS50132; RGS; 1.  
 DR Anti-Oncogene; Phosphorylation.  
 KW DOMAIN 81 200  
 FT DOMAIN 327 413 GSK-3B BINDING SITE (BY SIMILARITY).  
 FT DOMAIN 413 478 BETA-CATENIN BINDING SITE (BY  
 FT SIMILARITY).  
 FT POLY-HIS.  
 FT DOMAIN 469 476 DIX.  
 FT DOMAIN 758 840  
 FT CONFLICT 101 101 R -> K (IN REF. 2).  
 FT CONFLICT 474 474 H -> Y (IN REF. 2).  
 FT CONFLICT 484 484 S -> P (IN REF. 2).  
 FT CONFLICT 503 503 F -> S (IN REF. 2).  
 FT CONFLICT 603 603 G -> A (IN REF. 2).  
 SEQUENCE 840 AA: 92934 MW: A07D5EFB25DE7277 CRC64;

Query Match 4.9%; Score 8; DB 1; Length 840;  
 Best Local Similarity 100.0%; Pred. No. 3.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 ATASVST 17  
 |||||  
 DB 257 ATASVST 264  
 RESULT 6  
 AXIN2\_HUMAN  
 ID AXIN2\_HUMAN STANDARD; PRT; 843 AA.  
 AC 09V2T1: 09V84;  
 DT 16-OCT-2001 (rel. 40. Created)  
 DT 16-OCT-2001 (rel. 40. Last sequence update)  
 DT 16-OCT-2001 (rel. 40. Last annotation update)  
 DE Axin 2 (Axin inhibition protein 2) (conductin) (Axin-like protein)  
 DE (Ax11).  
 GN AXIN2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99168905; PubMed-10049590;  
 RA Mai M., Qian C., Yokomizo A., Smith D.I., Liu W.;  
 RT "Cloning of the human homolog of conductin (AXIN2), a gene mapping to  
 RT chromosome 17q23-q24.";  
 RL Genomics 55:341-344(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Lymphoblast;  
 RA Zhang T., Fagotto F., Hsu W., Zeng L., Gilbert D., Copeland N.G.,  
 RA Jenkins N.A., Warburton D., Costantini F.;  
 RT "Properties of mouse Axin2 and human AXIN2: chromosomal location,  
 RT expression pattern, interaction with Axin and effects on embryonic  
 RT axis formation.";  
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
 CC -I- FUNCTION: INHIBITOR OF THE WNT SIGNALING PATHWAY. DOWN-REGULATES  
 CC BETA-CATENIN. PROBABLY FACILITATE THE PHOSPHORYLATION OF BETA-  
 CC CATEININ AND APC BY GSK-3B (BY SIMILARITY).  
 CC -I- SUBUNIT: INTERACTS WITH GLYCOGEN SYNTHASE KINASE-3 BETA (GSK-3B)  
 CC AND BETA-CATENIN. THE INTERACTION BETWEEN AXIN AND BETA-CATENIN  
 CC OCCURS VIA THE ARMADILLO REPEATS CONTAINED IN BETA-CATENIN.  
 CC TERNARY COMPLEX (BY SIMILARITY).  
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND LYMPHOBLAST.  
 CC -I- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B AND DEPHOSPHORYLATED BY  
 CC PP2A (BY SIMILARITY).  
 CC -I- SIMILARITY: CONTAINS 1 RGS DOMAIN.  
 CC -I- SIMILARITY: CONTAINS 1 DIX DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF078165; AAD20976.1; -;  
 DR EMBL: AF205888; AAF22799.1; -;  
 DR HSSP: P49799; IAGR.  
 DR MIM: 604025; -;  
 DR InterPro: IPR001158; DIX.  
 DR InterPro: IPR000342; RGS.  
 DR Pfam: PF00778; DIX; 1.  
 DR Pfam: PF00615; RGS; 1.  
 DR PRINTS: PR01301; RGS-PROTEIN.  
 DR PRODOM: PD001580; RGS; 1.  
 DR Prodom: PD003639; DIX; 1.  
 DR SMART: SM00021; DAX; 1.  
 DR SMART: SM00315; RGS; 1.  
 DR PROSITE: PS50132; RGS; 1.  
 KW Developmental protein; Phosphorylation.

```

FT DOMAIN 81 200 RGS.
FT DOMAIN 327 413 GSK-3B BINDING SITE (BY SIMILARITY).
FT DOMAIN 413 476 BETA-CATENIN BINDING SITE (BY
FT DOMAIN 469 474 SIMILARITY).
FT DOMAIN 761 843 POLY-HIS.
FT CONFLICT 37 62 DIX.
FT CONFLICT 346 346 O -> R (IN REF. 2).
FT CONFLICT 572 636 MISSING (IN REF. 2).
FT CONFLICT 687 687 P -> S (IN REF. 2).
FT CONFLICT 696 696 Q -> H (IN REF. 2).
SQ SEQUENCE 843 AA; 93557 MW; F7B62BED6AB464D CRC64;

```

```

Query Match 4.9%; Score 8; DB 1; Length 843;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 10 ATASVRSF 17
Db 257 ATASVRSF 264

```

```

RESULT 7
ID IOVO_CIRAE STANDARD; PRT: 54 AA.
AC P05579;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ovomucoid (Fragment).
OS Circus aeruginosus (Marsh harrier).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Falconiformes; Accipitridae;
OC Accipitridae; Circus.
OX NCBI_TaxID=8964;
RN [1]
RP SEQUENCE.
RX MEDLINE=87157615; PubMed=3828298;
RA Laskowski M. Jr., Kato I., Ardelt W., Cook J., Denton A., Emple M.W.,
RA Kohr W.J., Park S.J., Parks R., Schatzley B.L., Schoenberger O.L.,
RA Tashiro M., Vichot G., Whalley H.E., Wleczorek A., Wleczorek M.;
RT "Ovomucoid third domains from 100 avian species: isolation,
RT sequences, and hypervariability of enzyme-inhibitor contact
RT residues."
RL Biochemistry 26:202-221(1987).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: AVIAN OVOMUCOID CONSIST OF THREE HOMOLOGOUS, TANDEM KAZAL
CC FAMILY INHIBITORY DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 KAZAL-LIKE DOMAINS.
CC PIR: B31443; B31443.
DR HSSP: P05586; 4OVO.
DR InterPro: IPR002350; Kazal.
DR Pfam: PF00050; Kazal; 1.
DR SMART: SM00280; KAZAL; 1.
DR PROSITE: PS00282; KAZAL; 1.
KW Serine protease inhibitor; Glycoprotein; Repeat; Egg white.
FT NON_TER 1
FT DOMAIN 1 54 KAZAL-LIKE 3.
FT ACT_SITE 16 17 REACTIVE BOND 3.
FT DISULFID 14 33
FT DISULFID 22 34
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .).
FT NON_TER 54
SQ SEQUENCE 54 AA; 5876 MW; 416C5E9A6DEB3AD1 CRC64;

```

```

Query Match 4.3%; Score 7; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 49 NGTSLIS 55
Db 43 NGTSLIS 49

```

```

RESULT 8
ID IOVO_OPIHO STANDARD; PRT: 54 AA.
AC P52249;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ovomucoid (Fragment).
OS Opisthocomus hoazin (Hoatzin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Opisthocomiformes; Opisthocomidae;
OC Opisthocomus.
OX NCBI_TaxID=30419;
RN [1]
RP SEQUENCE.
RX MEDLINE=94072078; PubMed=8251062;
RA Apostol I., Giletto A., Komiya T., Zhang W., Laskowski M. Jr.;
RT "Amino acid sequences of ovomucoid third domains from 27 additional
RT species of birds."
RL J. Protein Chem. 12:419-433(1993).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: AVIAN OVOMUCOID CONSIST OF THREE HOMOLOGOUS, TANDEM KAZAL
CC FAMILY INHIBITORY DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 KAZAL-LIKE DOMAINS.
CC HSSP: P05586; 4OVO.
DR InterPro: IPR002350; Kazal.
DR Pfam: PF00050; Kazal; 1.
DR SMART: SM00280; KAZAL; 1.
DR PROSITE: PS00282; KAZAL; 1.
KW Serine protease inhibitor; Glycoprotein; Repeat; Egg white.
FT NON_TER 1
FT DOMAIN 1 54 KAZAL-LIKE 3.
FT ACT_SITE 16 17 REACTIVE BOND 3.
FT DISULFID 14 33
FT DISULFID 22 34
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .).
FT NON_TER 54
SQ SEQUENCE 54 AA; 5828 MW; 95B10804D446583D CRC64;

```

```

Query Match 4.3%; Score 7; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 49 NGTSLIS 55
Db 43 NGTSLIS 49

```

```

RESULT 9
ID PER_DROAN STANDARD; PRT: 125 AA.
AC 003293;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Period circadian protein (Fragment).
GN PER.
OS Drosophila ananassae (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7217;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93196482; PubMed=8450754;
RA Peixoto A.A., Campesan S., Costa R.H., Kyriacou C.P.;

```

```

RT "Molecular evolution of a repetitive region within the per gene of
RT Drosophila."
RL Mol. Biol. Evol. 10:127-139(1993).
CC -1- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS: AN INCREASE IN
CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
CC RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE
CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
CC TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING
CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
CC -1- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN
CC TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
CC TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
CC -1- DOMAIN: THE RUN OF GLY-THR IS IMPLICATED IN THE MAINTENANCE OF
CC CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT
CC LEADS TO A SHORTENING OF THE COURSHIP SONG CYCLE PERIOD, AND THUS
CC COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC
CC MATING BEHAVIOR (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
CC PER-TIM (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L06335; AAA28758.1; -.
DR FLYBASE: FBgn0012142; DanaPer.
KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
FT NON_TER 1 1
FT DOMAIN 30 77 G-T REPEATS.
FT DOMAIN 96 102 POLY-THR.
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 11118 MW; C733DD3586458A6F CRC64;

Query Match 4.3%; Score 7; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TTTAATA 12
DB 105 TTTAATA 111

RESULT 10
HN3G_MOUSE STANDARD; PRT; 353 AA.
AC P35584;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hepatocyte nuclear factor 3-gamma (HNF-3G).
GN HNF3G OR FOXA3 OR TCF3G OR TCF-3G.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94307723; PubMed=8034310;
RA Kaetner K., Hiemisch H., Luckow B., Schuetz G.;
RT "The Hnf-3 gene family of transcription factors in mice: gene
RT structure, cDNA sequence, and mRNA distribution.";
RL Genomics 20:377-385(1994).
CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LIVER GENES
CC SUCH AS AFP, ALBUMIN, TYROSINE AMINOTRANSFERASE, PEPC, ETC.
CC INTERACTS WITH THE CIS-ACTING REGULATORY REGIONS OF THESE GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: RESTRICTED MAINLY TO ENDODERM-DERIVED TISSUES
CC (LUNG, LIVER, STOMACH, AND SMALL INTESTINE), ALSO PRESENT
CC ADDITIONALLY IN OVARY, TESTIS, HEART, AND ADIPOSE TISSUE, BUT
CC MISSING FROM LUNG.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION PEAKS AROUND DAY 15.5 OF
CC GESTATION.
CC -----
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74938; CAA52892.1; -.
DR PIR: S37186; S37186.
DR HSSP: 063245; 2HFH.
DR TRANSFAC: T02345; -.
DR MGD: MGI:1347477; Foxa3.
DR InterPro: IPR001766; Fork head.
DR Pfam: PF00250; Fork head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR SMART: SM00339; FH; 1.
DR PROSITE: PS00657; FORK_HEAD_1; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
DR DNA-binding: Nuclear protein; Transcription regulation; Activator.
KW DNA BIND 118 209 FORK-HEAD
FT SEQUENCE 353 AA; 37601 MW; 28F060A8E944D5B9 CRC64;

Query Match 4.3%; Score 7; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TTTAATA 12
DB 243 TTTAATA 249

RESULT 11
HN3G_RAT STANDARD; PRT; 354 AA.
AC P32183;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hepatocyte nuclear factor 3-gamma (HNF-3G).
GN HNF3G OR TCF3G OR TCF-3G.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91160974; PubMed=1672118;
RA Lai E., Prezioso V.R., Tao W.F., Chen W.S., Darnell J.E. Jr.;
RT "Hepatocyte nuclear factor 3 alpha belongs to a gene family in
RT mammals that is homologous to the Drosophila homeotic gene fork

```

```

RT head.
RL Genes Dev. 5:416-427(1991).
RN [2]
RP STRUCTURE BY NMR OF 107-223.
RX MEDLINE-93323996; PubMed-6332212.
RA Clark K.L., Halsey E.D., Lai E., Burley S.K.;
RT "Co-crystal structure of the HNF-3/Fork head DNA-recognition motif
RT resembles histone H5."
RL Nature 364:412-420(1993).
CC -1- FUNCTION: TRANSCRIPTION ACTIVATOR FOR A NUMBER OF LIVER GENES
CC SUCH AS AFP, ALBUMIN, TYROSINE AMINOTRANSFERASE, PERCK, ETC.
CC INTERACTS WITH THE CIS-ACTING REGULATORY REGIONS OF THESE GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Liver.
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L09648; AAA41339.1; -
DR PIR: S35090; S35090.
DR HSSP: Q63245; 2HFH.
DR TRANSFAC: T01050; -
DR InterPro: IPR001766; Fork_head.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR SMART: SM00339; FH; 1.
DR PROSITE: PS00657; FORK_HEAD_1; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
DR DNA-binding: Nuclear protein; Transcription regulation; Activator.
FT DNA BIND 118 209 FORK-HEAD.
SQ SEQUENCE 354 AA; 37652 MW; 318B01ECCE/C365C CRC64;

Query Match 4.3%; Score 7; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTTAATA 12
Db 243 TTTAATA 249

RESULT 12
AMP_C_PSYIM STANDARD; PRT; 401 AA.
ID AMP_C_PSYIM
AC 005465;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Beta-lactamase precursor (EC 3.5.2.6) (Cephalosporinase).
GN AMP_C OR BLA.
OS Psychrobacter immobilis.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Psychrobacter.
OX NCBI_TaxID:498;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 40-47, AND CHARACTERIZATION.
RX STRAIN-A5;
RX MEDLINE-97217442; PubMed-9063463;
RA Feller G., Zekhnin Z., Lamotte-Brasseur J., Gerday C.H.;
RT "Enzymes from cold-adapted microorganisms. The class C beta-lactamase
RT from the antarctic psychrophile Psychrobacter immobilis A5";
RL Eur. J. Biochem. 244:186-191(1997).
CC -1- FUNCTION: THIS PROTEIN IS A SERINE BETA-LACTAMASE WITH A SUBSTRATE
CC SPECIFICITY FOR CEPHALOSPORINS.
CC -1- CATALYTIC ACTIVITY: A beta-lactam + H(2)O = a substituted beta-
```

```

CC -1- amino acid.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE CLASS-C BETA-LACTAMASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83586; CA58569.1; -
DR HSSP: P05364; 2BLT.
DR InterPro: IPR001466; Beta_Lactam.
DR InterPro: IPR001586; Beta_Lactam_C.
DR Pfam: PF00144; beta-lactamase; 2.
DR PROSITE: PS00336; BETA_LACTAMASE_C; 1.
DR Hydrolyase; Antibiotic resistance; Signal.
FT SIGNAL 1 39
FT CHAIN 40 401 BETA-LACTAMASE.
FT ACT_SITE 102 102 BY SIMILARITY.
FT BINDING 353 355 SUBSTRATE (BY SIMILARITY).
SQ SEQUENCE 401 AA; 44450 MW; 93F0DB278EA8E043 CRC64;

Query Match 4.3%; Score 7; DB 1; Length 401;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 LNTGLSL 54
Db 117 LNTGLSL 123

RESULT 13
KAS2_STRCM STANDARD; PRT; 402 AA.
ID KAS2_STRCM
AC P41176;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE putative polyketide beta-ketoacyl synthase 2 (EC 2.3.1.-) (ORF 2).
OS Streptomyces cinamonensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID:1900;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-A3823.5;
RX MEDLINE-92374994; PubMed-1508151;
RA Arrowsmith T.J., Malpartida F., Sherman D.H., Birch A., Hopwood D.A.,
RA Robinson J.A.;
RT "Characterisation of acti-homologous DNA encoding polyketide synthase
RT genes from the menshyn producer Streptomyces cinamonensis."
RL Mol. Gen. Genet. 234:254-264(1992).
CC -1- PATHWAY: BIOSYNTHESIS OF THE POLYKETIDE ANTIBIOTIC MONENCTIN.
CC -1- MISCELLANEOUS: THIS PUTATIVE KETOACYL SYNTHASE LACKS THE ACTIVE
CC SITE CYSTEINE.
CC -1- SIMILARITY: BELONGS TO THE BETA-KETOACYL-ACP SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z11511; CA477597.1; -
DR PIR: S25077; S25077.
DR InterPro: IPR000794; Ketoacyl-synt.
DR Pfam: PF00109; ketoacyl-synt. 1.
```

DR Pfam: PF02801; ketoacyl-synt\_C.1.  
 KW Antibiotic biosynthesis; Transferase; Acyltransferase.  
 SQ SEQUENCE 402 AA: 41270 MW: 348657D10201C214 CRC64;

Query Match 4.3%; Score 7; DB 1; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 78 EHLPRRL 84  
 |||||  
 DB 59 EHLPRRL 65

RESULT 14  
 ID PGTI\_HUMAN STANDARD; PRT: 500 AA.  
 AC 016647;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Prostacyclin synthase (EC 5.3.99.4) (Prostaglandin I2 synthase).  
 GN PTGIS OR CYP8 OR CYP8A1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Arteria;  
 RX MEDLINE=94242046; PubMed=8185632;  
 RA Miyata A., Hara S., Yokoyama C., Inoue H., Ullrich V., Tanabe T.;  
 RT "Molecular cloning and expression of human prostacyclin synthase";  
 RL Biochem. Biophys. Res. Commun. 200:1728-1734(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. AND VARIANTS LEU-38; ARG-118 AND SER-379.  
 RX MEDLINE=21177061; PubMed=11281454;  
 RA Chevalier D., Cauffiez C., Bernard C., Lo-Guidice J.-M., Allorge D.,  
 RA Pazio F., Ferrari N., Libersa C., Lhermitte M., D'Halluin J.C.,  
 RA Broly F.;  
 RT "Characterization of new mutations in the coding sequence and  
 5'-untranslated region of the human prostacyclin synthase gene  
 (CYP8A1)";  
 RL Hum. Genet. 108:148-155(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Delouis P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggsley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leivasalho M.H., Laversha M., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurtry A.A.,  
 RA Millne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showken R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20";  
 RL Nature 414:865-871(2001).

CC - FUNCTION: CATALYZES THE ISOMERIZATION OF PROSTAGLANDIN H2 TO  
 CC PROSTACYCLIN (= PROSTAGLANDIN I2).  
 CC - CATALYTIC ACTIVITY: (5*Z*,13*E*)-(1*S*)-9- $\alpha$ ,11- $\alpha$ -epidioxo-15-  
 CC hydroxyprosta-5,13-dienoate = (5*Z*,13*E*)-(1*S*)-6,9- $\alpha$ -epoxy-11-  
 CC  $\alpha$ ,15-dihydroxyprosta-5,13-dienoate.  
 CC - SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE. MAY BE  
 CC ANCHORED TO THE MEMBRANE VIA A SINGLE TRANSMEMBRANE DOMAIN.  
 CC - TISSUE SPECIFICITY: WIDELY EXPRESSED; PARTICULARLY ABUNDANT IN  
 CC OVARY, HEART, SKELETAL MUSCLE, LUNG, AND PROSTATE.  
 CC - SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC - DATABASE: NAME=cytochrome p450 Allele Nomenclature Committee;  
 CC NOTE=CYP8A1 alleles;  
 CC WWW="http://www.imm.ki.se/cypalleles/cyp8a1.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D38145; BAA07343.1; -;  
 DR EMBL: AF287048; AAG31781.1; -;  
 DR EMBL: AF287049; AAG31782.1; -;  
 DR EMBL: AL118525; CAC14162.1; -;  
 DR MIM: 601699; -;  
 DR InterPro: IPR001128; Cyt\_P450.  
 DR Pfam: PF00067; P450.2.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; FALSE\_NEG.  
 KW Isomerase; Prostaglandin biosynthesis; Heme; Transmembrane;  
 KW Endoplasmic reticulum; Polymorphism.  
 FT BINDING 1 20 POTENTIAL.  
 FT VARIANT 38 38 HEME (BY SIMILARITY).  
 FT VARIANT 118 118 P -> L (IN CYP8A1\*2).  
 FT VARIANT 118 118 S -> R (IN CYP8A1\*3).  
 FT VARIANT 379 379 /FTID=VAR\_010916.  
 FT VARIANT 379 379 R -> S (IN CYP8A1\*4).  
 FT VARIANT 379 379 /FTID=VAR\_010917.  
 SQ SEQUENCE 500 AA: 57103 MW: 39595442BFC0B625 CRC64;  
 Query Match 4.3%; Score 7; DB 1; Length 500;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 PVSQTTT 8  
 |||||  
 DB 323 PVSQTTT 329  
 RESULT 15  
 ID EFG1\_CANAL STANDARD; PRT: 552 AA.  
 AC P43064;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Enhanced filamentous growth protein.  
 GN EFG1 OR EFG.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 10231;  
 RX MEDLINE=97299875; PubMed=9155024;  
 RA Stoldt V.R., Sonneborn A., Leuker C.E., Ernst J.F.;  
 RT "Efg1p, an essential regulator of morphogenesis of the human pathogen  
 Candida albicans, is a member of a conserved class of bHLH proteins



```

RT regulating morphogenetic processes in fungi."
RL EMBL J. 16:1982-1991(1997).
CC -I- FUNCTION: PUTATIVE TRANSCRIPTION FACTOR THAT STIMULATES
CC PSEUDOHYPHAL MORPHOGENESIS.
CC -I- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -I- SIMILARITY: BELONGS TO THE EFG1/PHD1/STU4 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z32687; CAAB3640.1; -.
DR InterPro; IPR003163; Yeast_DNA_bind.
DR Pfam; PF02292; Yeast_DNA_bind; 1.
KW Transcription regulation; Nuclear protein; DNA-binding.
FT DOMAIN 34 37 POLY-GLN.
FT DOMAIN 46 50 POLY-GLN.
FT DOMAIN 83 93 POLY-GLN.
FT DOMAIN 315 326 POLY-ALA.
FT DOMAIN 332 338 POLY-PRO.
FT DOMAIN 438 446 POLY-GLN.
FT DOMAIN 470 473 POLY-GLN.
FT DOMAIN 483 493 POLY-GLN.
FT DOMAIN 496 499 POLY-GLN.
SQ SEQUENCE 552 AA; 59964 MW; F94FD94FC2E06EB7 CRC64;

```

```

Query Match 4.3%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 TTTAATA 12
   |||||
Db 343 TTTAATA 349

```

Search completed: October 9, 2002, 02:52:57  
 Job time: 3938 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 9, 2002, 00:57:49 ; Search time 38.03 Seconds

(without alignments)  
414.374 Million cell updates/sec

Title: US-09-786-130-1

Perfect score: 164

Sequence: 1 TPVSGTTATATASVSTKDP.....TLPPQGEALPSSHSPPQOG 164

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: PIR-71:\*\*

1: PIR1:\*\*

2: PIR2:\*\*

3: PIR3:\*\*

4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	11.6	235	2 T30656	hypothetical prote
2	8	4.9	290	2 T50434	hypothetical prote
3	8	4.9	368	2 S74377	riboflavin biunct
4	8	4.9	443	2 E64725	yeau protein - Esc
5	8	4.9	443	2 H90634	probable transport
6	8	4.9	443	2 H85485	probable transport
7	8	4.9	520	2 T04136	cell death suppress
8	8	4.9	838	2 T08423	Axin homolog Axil
9	7	4.3	54	2 B31443	ovomucoid, third d
10	7	4.3	54	2 H61588	hypomucoid (PSTI-ty
11	7	4.3	147	2 T49624	hypothetical prote
12	7	4.3	176	2 F83531	hypothetical prote
13	7	4.3	246	2 C70527	hypothetical prote
14	7	4.3	310	2 A70421	fructose-1,6-bisph
15	7	4.3	330	2 S37595	mucin JvU10 - huma
16	7	4.3	339	2 A72609	probable deblockin
17	7	4.3	353	2 C54258	transcription fact
18	7	4.3	354	1 S35090	transcription fact
19	7	4.3	402	1 S25077	momenin polyketid
20	7	4.3	500	1 JC2231	prostaglandin-I sy
21	7	4.3	500	2 E83882	alpha-L-arabino fur
22	7	4.3	551	2 T01832	hypothetical prote
23	7	4.3	552	2 S49338	transcription fact
24	7	4.3	594	2 D97431	probable alpha-glu
25	7	4.3	600	2 T46880	succinate dehydrog
26	7	4.3	613	2 D97676	succinate dehydrog
27	7	4.3	613	2 AB2901	succinate dehydrog
28	7	4.3	630	2 AD3272	succinate dehydrog
29	7	4.3	702	2 A69140	ATP-dependent heli

30	7	4.3	710	2 S66313	phenylalanine ammo
31	7	4.3	766	2 S45344	trp1 like enhancer
32	7	4.3	779	2 T21021	hypothetical prote
33	7	4.3	891	2 T29561	hypothetical prote
34	7	4.3	1017	2 I37465	HIRA protein - hum
35	7	4.3	1042	2 S76045	hypothetical prote
36	7	4.3	1069	2 S27922	nuclear antigen EB
37	7	4.3	1213	2 T51032	hypothetical prote
38	7	4.3	1348	2 S27812	probable epidermal
39	7	4.3	1348	2 A43917	probable epidermal
40	7	4.3	1388	2 T00063	hypothetical prote
41	7	4.3	1617	2 B86483	protein F5J5.15 (I
42	7	4.3	5126	2 S40450	ryanodine receptor
43	6	3.7	22	2 G30608	ig kappa chain V-I
44	6	3.7	26	2 G30608	ig kappa chain V-I
45	6	3.7	26	2 D30607	ig kappa chain V-I

## ALIGNMENTS

RESULT 1  
T30656  
hypothetical protein 54L - Molluscum contagiosum virus 1  
N:Alternate names: MC054L  
C:Species: Molluscum contagiosum virus 1  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 11-May-2000  
C:Accession: T30656  
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
Science 273, 813-816, 1996  
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host  
A:Reference number: Z20876; M01D:96325459  
A:Accession: T30656  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-235 <SEN>  
A:Cross-references: EMBL:U60315; P1DN:AMC5182.1  
C:Genetics:  
A:Note: MC054L

Query Match 11.6%; Score 19; DB 2; Length 235;  
Best Local Similarity 100.0%; Pred. No. 9.9e-12;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 HSTNFCVLDPEDQVGRH 132  
Db 113 HSTNFCVLDPEDQVGRH 131

RESULT 2  
T50434  
hypothetical protein SPCC4B3.14 [imported] - fission yeast (Schizosaccharomyces pombe  
C:Species: Schizosaccharomyces pombe  
C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000  
C:Accession: T50434  
R:Seeger, K.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, November 1999  
A:Reference number: Z25033  
A:Accession: T50434  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-290 <SEN>  
A:Cross-references: EMBL:AL132870; P1DN:CAB60688.1; GSPDB:GN00068; SPDB:SPCC4B3.14  
A:Experimental source: strain 972h(-); cosmid c4B3  
C:Genetics:  
A:Gene: SPDB:SPCC4B3.14  
A:Map position: 3

Query Match 4.9%; Score 8; DB 2; Length 290;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

F;23-41/domain: transmembrane #status predicted <IML>

c:\superfamily: yaad proteins

```
Query Match          4.9%; Score 8; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 LEOLTPPAL 113
|||||
Db 41 LEOLTPPAL 48

RESULT 7
T04136
cell death suppressor protein 11s1 - maize
N:Alternate names: lethal leaf-spot 1
C:Species: Zea mays (maize)
C>Date: 23-Apr-1999 #sequence_revision 20-Aug-1999 #text_change 21-Jul-2000
C:Accession: T04136; T04133
R:Gray, J.; Close, P. S.; Briggs, S. P.; Johal, G. S.
Cell 89, 25-31, 1997
A:Title: A novel suppressor of cell death in plants encoded by the 11s1 gene of maize.
A:Reference number: Z15232; MUID:97248483
A:Accession: T04136
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-505 <GRA1>
A:Cross-references: EMBL:U77346; NID:g1935910; PIDN:AAC49678.1; PID:g1935912
A:Accession: T04133
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 54-403; 'F', 405-443; 'HT', 446-497; 'V', 499-520 <GRA>
A:Cross-references: EMBL:U77345; NID:g1935908; PIDN:AAC49676.1; PID:g1935909
A:Experimental source: strain pa405 inbred line; leaf
C:Genetics:
A:Gene: 11s1
A:Map position: 15
A:introns: 123/3; 202/2; 244/3; 290/2; 343/3; 436/3
C:Superfamily: Rieske [2Fe-2S] homology
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein
F:105-153/Domain: Rieske [2Fe-2S] homology <RSK>
F:115,117,135,138/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status F

Query Match          4.9%; Score 8; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 PTQELAPS 155
|||||
Db 422 PTQELAPS 429

RESULT 8
T08423
Axin homolog Ax1l - rat
N:Alternate names: Ax1l
C:Species: Rattus norvegicus (Norway rat)
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jul-2000
C:Accession: T08423
R:Yamamoto, H.; Kishida, S.; Uochi, T.; Ikeda, S.; Koyama, S.; Asashima, M.; Kikuchi, A.
Mol. Cell. Biol. 18, 2867-2875, 1998
A:Title: Ax1l, a member of the Axin family, interacts with both glycogen synthase kinase
A:Reference number: Z16414; MUID:98226558
A:Accession: T08423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-838 <YAM>
A:Cross-references: EMBL:AF017757; NID:g3080758; PIDN:AAC40089.1; PID:g3080759
A:Note: interacts with GSK-3beta and beta-catenin
C:Keywords: phosphoprotein; signal transduction

Query Match          4.9%; Score 8; DB 2; Length 838;
Best Local Similarity 100.0%; Pred. No. 7.6;
```

```
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 ATASVRSST 17
|||||
Db 257 ATASVRSST 264

RESULT 9
B31443
ovomucoid, third domain - western marsh harrier (fragment)
C:Species: Circus aeruginosus (western marsh harrier)
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 07-Feb-1997
C:Accession: B31443
R:Laskowski Jr., M.; Kato, I.; Ardelt, W.; Cook, J.; Denton, A.; Empie, M.W.; Kohr, W.
A:Biochemistry 26, 202-221, 1987
A:Title: Ovomucoid third domains from 100 avian species: Isolation, sequences, and hy
A:Reference number: A90515; MUID:87157615
A:Accession: B31443
A:Molecule type: protein
A:Residues: 1-54 <LAS>
A:Note: the authors designate this sequence with the code OMMAH3
C:Superfamily: ovomucoid; Kazal proteinase inhibitor homology
C:Keywords: egg white; glycoprotein; serine proteinase inhibitor
F:4-54/Domain: Kazal proteinase inhibitor homology <KRP>
F:6-36,14-33,22-54/Disulfide bonds: #status predicted
F:43/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match          4.3%; Score 7; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 NGTISLS 55
|||||
Db 43 NGTISLS 49

RESULT 10
H61588
ovomucoid (PSII-type proteinase inhibitor), third domain - hoatzin
C:Species: Opisthocomus hoazin (hoatzin)
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 17-Mar-1999
C:Accession: H61588; A44435
R:Apostol, I.; Giletti, A.; Komiya, T.; Zhang, W.; Laskowski Jr., M.
J. Protein Chem. 12, 419-433, 1993
A:Title: Amino acid sequences of ovomucoid third domains from 27 additional species o
A:Reference number: A61587; MUID:94072078
A:Accession: H61588
A:Molecule type: protein
A:Residues: 1-54 <APO>
C:Superfamily: ovomucoid; Kazal proteinase inhibitor homology
C:Keywords: serine proteinase inhibitor
F:4-54/Domain: Kazal proteinase inhibitor homology <KRP>
F:6-36,14-33,22-54/Disulfide bonds: #status predicted

Query Match          4.3%; Score 7; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 NGTISLS 55
|||||
Db 43 NGTISLS 49

RESULT 11
T49624
hypothetical protein B5022.60 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C:Accession: T49624
R:Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
```

submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022

A:Accession: T49624

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-147 <SCH>

A:Cross-references: EMBL:AL355932; GSPDB:GN00116; NCSP:B5022.60

A:Experimental source: BAC clone B5022; strain OR74A

C:Genetics:

A:Gene: NCSP:B5022.60

A:Map position: 6

C:Superfamily: Neurospora crassa hypothetical protein B5022.60

#### Query Match

Best Local Similarity 4.3%; Score 7; DB 2; Length 147;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CSRPNF 65

DB 30 CSRPNF 36

#### RESULT 12

F83531  
hypothetical protein PA0907 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: F83531

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lam,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: AB2950; MUID:20437337

A:Accession: F83531

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-176 <STO>

A:Cross-references: GB:AE004525; GB:AE004091; NID:g9946805; PIDN:AAG04296.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0907

#### Query Match

Best Local Similarity 4.3%; Score 7; DB 2; Length 176;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SFIEHLP 81

DB 40 SFIEHLP 46

#### RESULT 13

C70527  
hypothetical protein RV0330c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: C70527

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98293987

A:Accession: C70527

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-246 <COI>

A:Cross-references: GB:Z96800; GB:AL123456; NID:g3261800; PIDN:CAB09578.1; PID:e321793;

A:Experimental source: strain H37RV

C:Genetics:  
A:Gene: RV0330c

#### Query Match

Best Local Similarity 4.3%; Score 7; DB 2; Length 246;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 LRATLPP 148

DB 106 LRATLPP 112

#### RESULT 14

A70421  
fructose-1,6-bisphosphate aldolase class II - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 22-Jun-1999

C:Accession: A70421

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666

A:Accession: A70421

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-310 <AOE>

A:Cross-references: GB:AE000737; NID:g2983782; PIDN:AAC07345.1; PID:g2983787; GB:AE00

A:Experimental source: strain VF5

C:Genetics:

A:Gene: fba

#### Query Match

Best Local Similarity 4.3%; Score 7; DB 2; Length 310;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 VLVDPEQ 127

DB 151 VLVDPEQ 157

#### RESULT 15

S37595  
mucin JUI10 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999

C:Accession: S37595; S35048

R:Aubert, J.

submitted to the EMBL Data Library, September 1993

A:Reference number: S37593

A:Accession: S37595

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-330 <ADU>

A:Cross-references: EMBL:X74956; NID:g407052; PIDN:CAA52911.1; PID:g407053

R:Dufores, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperrat, V.; Laine, A.; Van-Seuning

Biochem. J. 293, 329-337, 1993

A:Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic altern

A:Reference number: S35047; MUID:93343858

A:Accession: S35048

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-330 <DUF>

#### Query Match

Best Local Similarity 4.3%; Score 7; DB 2; Length 330;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTTAATA 12

Db 34 |||||  
TTTATA 40

Search completed: October 9, 2002, 02:25:38  
Job time: 5269 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2002, 22:31:48 ; Search time 19.91 Seconds  
(without alignments)  
201.195 Million cell updates/sec

Title: US-09-786-130-1  
Perfect score: 164  
Sequence: 1 TPVSQTTATATASVRSSTKDP.....TLPTQALPSSHPQOOG 164

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.9	520	4	US-08-810-009-2
2	7	4.3	50	1	US-08-063-551A-2
3	7	4.3	50	3	US-08-926-842B-50
4	7	4.3	116	1	US-08-478-039-103
5	7	4.3	116	1	US-08-476-349A-103
6	7	4.3	275	2	US-08-578-709-13
7	7	4.3	385	4	US-09-036-987A-19
8	7	4.3	385	4	US-09-370-700-19
9	7	4.3	403	4	US-08-861-774E-27
10	7	4.3	406	4	US-08-861-774E-25
11	7	4.3	407	4	US-08-861-774E-28
12	7	4.3	500	2	US-08-578-709-15
13	6	3.7	10	4	US-09-238-448-1
14	6	3.7	10	4	US-09-238-448-5
15	6	3.7	14	2	US-07-876-941A-36
16	6	3.7	49	3	US-08-926-842B-49
17	6	3.7	57	1	US-08-162-102C-38
18	6	3.7	84	3	US-08-906-769-99
19	6	3.7	84	3	US-08-906-616-99
20	6	3.7	84	4	US-08-817-795-99
21	6	3.7	84	4	US-08-639-075A-99
22	6	3.7	84	4	US-09-012-431-99
23	6	3.7	84	4	US-09-012-692-99
24	6	3.7	84	4	US-08-906-613-99
25	6	3.7	84	5	PCT-US95-14442A-99
26	6	3.7	93	3	US-08-894-699-63
27	6	3.7	93	4	US-09-444-410-63

28	6	3.7	94	3	US-08-894-699-62	Sequence 62, App1
29	6	3.7	94	3	US-08-894-699-64	Sequence 64, App1
30	6	3.7	94	4	US-09-444-410-62	Sequence 62, App1
31	6	3.7	94	4	US-09-444-410-64	Sequence 64, App1
32	6	3.7	96	4	US-09-228-986-98	Sequence 98, App1
33	6	3.7	97	3	US-08-894-699-26	Sequence 26, App1
34	6	3.7	97	3	US-08-894-699-27	Sequence 27, App1
35	6	3.7	97	3	US-08-894-699-30	Sequence 30, App1
36	6	3.7	97	4	US-09-444-410-26	Sequence 26, App1
37	6	3.7	97	4	US-09-444-410-27	Sequence 27, App1
38	6	3.7	97	4	US-09-444-410-30	Sequence 30, App1
39	6	3.7	99	3	US-08-894-699-31	Sequence 31, App1
40	6	3.7	99	3	US-08-894-699-34	Sequence 34, App1
41	6	3.7	99	4	US-09-444-410-31	Sequence 31, App1
42	6	3.7	99	4	US-09-444-410-34	Sequence 34, App1
43	6	3.7	104	1	US-08-276-852-94	Sequence 94, App1
44	6	3.7	104	1	US-08-276-852-100	Sequence 100, App
45	6	3.7	104	1	US-08-899-575-94	Sequence 94, App1

#### ALIGNMENTS

```
RESULT 1
US-08-810-009-2
; Sequence 2, Application US/08810009
; Patient No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; City: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spivull, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-2

Query Match 4.9%; Score 8; DB 4; Length 520;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 148 PTOFALPS 155
|||||||
```

Db 422 PROEALPS 429

## RESULT 2

US-08-063-551A-2

; Sequence 2, Application US/08063551A  
; Patent No. 5434071

## ; GENERAL INFORMATION:

; APPLICANT: ROSENBERG, Eugene

; APPLICANT: SROHAM, Yuval

; TITLE OF INVENTION: Preparation exhibiting enzymatic

; TITLE OF INVENTION: delignification activity, a method of

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BACON &amp; THOMAS

; STREET: 625 Slaters Lane - 4th Floor

; CITY: Alexandria

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 1.2mb storage

; COMPUTER: IBM AT COMPATIBLE

; OPERATING SYSTEM: MS-DOS 5.0

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/063,551A

; FILING DATE: 19930518

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 854645

; FILING DATE: 02-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Richard E. Fichter

; REGISTRATION NUMBER: 26,382

; REFERENCE/DOCKET NUMBER: REF/Rosenberg/551

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-683-0500

; TELEFAX: 703-683-1080

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 50 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: N-terminal fragment

US-08-063-551A-2

OY 74 GSFIEHL 80  
Db 24 GSFIEHL 30Query Match 4.3%; Score 7; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 3

US-08-926-842B-50

; Sequence 50, Application US/08926842B

; Patent No. 6030807

## ; GENERAL INFORMATION:

; APPLICANT: Sa-No. 6030807ueira, Isabel

; APPLICANT: de Lencastre, Herminda

; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE

; NUMBER OF SEQUENCES: 64

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber &amp; Jackson

; STREET: 411 Hackensack Avenue

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/926,842B

; FILING DATE: 10-SEP-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-089 N

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201 487-5800

; TELEFAX: 201 343-1684

; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 50:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 50 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

US-08-926-842B-50

OY 74 GSFIEHL 80  
Db 24 GSFIEHL 30Query Match 4.3%; Score 7; DB 3; Length 50;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 4

US-08-478-039-103

; Sequence 103, Application US/08478039

; Patent No. 5681722

## ; GENERAL INFORMATION:

; APPLICANT: Newman, Roland A.

; APPLICANT: Raab, Ronald W.

; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER &amp; MATHIS

; STREET: 699 Prince St.

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-1404

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/478,039

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/379,072

; FILING DATE: 25-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/912,292

; FILING DATE: 10-JUL-1992

; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/856,281  
FILING DATE: 23-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/735,064  
FILING DATE: 25-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin Esq., Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-160  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Monkey  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: #9  
US-08-478-039-103

Query Match 4.3%; Score 7; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 LSLSCVA 58  
|||||||  
DB 18 LSLSCVA 24

RESULT 5  
US-08-476-349A-103  
Sequence 103, Application US/08476349A  
Patent No. 5750105  
GENERAL INFORMATION:  
APPLICANT: Newman, Roland A.  
APPLICANT: Hanna, Nabil  
APPLICANT: Raab, Ronald W.  
TITLE OF INVENTION: Recombinant Antibodies for Human Therapy  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince St.  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,349A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/379,072  
FILING DATE: 25-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/912,292  
FILING DATE: 10-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/856,281  
FILING DATE: 23-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/735,064  
FILING DATE: 25-JUL-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Teskin Esq., Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-161  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Monkey  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: #9  
US-08-476-349A-103

Query Match 4.3%; Score 7; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 LSLSCVA 58  
|||||||  
DB 18 LSLSCVA 24

RESULT 6  
US-08-578-709-13  
Sequence 13, Application US/08578709  
Patent No. 5814509  
GENERAL INFORMATION:  
APPLICANT: TANABE, Tadashi  
TITLE OF INVENTION: PROSTACYCLIN SYNTHASE DERIVED FROM HUMAN  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/578,709  
FILING DATE: 28-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/00838  
FILING DATE: 27-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 114316/1994  
FILING DATE: 28-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gubinsky, Louis  
REGISTRATION NUMBER: 24,835  
REFERENCE/DOCKET NUMBER: Q40439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 275 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-578-709-13

Query Match 4.3%; Score 7; DB 2; Length 275;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVSQTTT 8  
|||||  
DB 98 PVSQTTT 104

RESULT 7

US-09-036-987A-19  
; Sequence 19, Application US/09036987A  
; Patent No. 6143526  
; GENERAL INFORMATION:  
; APPLICANT: Baltz, Richard H.  
; APPLICANT: Broughton, Mary C.  
; APPLICANT: Crawford, Kathryn P.  
; APPLICANT: Madhuri, Krishnamurthy  
; APPLICANT: Merlo, Donald J.  
; APPLICANT: Turner, Jan R.  
; APPLICANT: Waldron, Clive  
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Dow Agrosciences LLC Patent Department  
; STREET: 9330 Zionsville Road  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/036,987A  
; FILING DATE: 09-MAR-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stuart, Donald R.  
; REGISTRATION NUMBER: 28,479  
; REFERENCE/DOCKET NUMBER: 50,608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317)337-4816  
; TELEFAX: (317)337-4847  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-036-987A-19

Query Match 4.3%; Score 7; DB 4; Length 385;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 VDEQVY 129  
|||||  
DB 45 VDEQVY 51

RESULT 8

US-09-370-700-19  
; Sequence 19, Application US/09370700  
; Patent No. 6274350

; GENERAL INFORMATION:  
; APPLICANT: Baltz, Richard H.  
; APPLICANT: Broughton, Mary C.  
; APPLICANT: Crawford, Kathryn P.  
; APPLICANT: Madhuri, Krishnamurthy  
; APPLICANT: Treadway, Patti J.  
; APPLICANT: Turner, Jan R.  
; APPLICANT: Waldron, Clive  
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
; FILE REFERENCE: 50489 DIV1  
; CURRENT APPLICATION NUMBER: US/09/370,700  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/36987  
; EARLIER FILING DATE: 1998-03-09  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO: 19  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Saccharopolyspora spinosa  
US-09-370-700-19

Query Match 4.3%; Score 7; DB 4; Length 385;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 VDEQVY 129  
|||||  
DB 45 VDEQVY 51

RESULT 9

US-08-861-774E-27  
; Sequence 27, Application US/08861774E  
; Patent No. 6297007  
; GENERAL INFORMATION:  
; APPLICANT: Waters, Barbara  
; APPLICANT: Miao, Vivian  
; APPLICANT: Ho, Yap  
; APPLICANT: Tong, Seow  
; TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR  
; FILE REFERENCE: 9993-006  
; CURRENT APPLICATION NUMBER: US/08/861,774E  
; CURRENT FILING DATE: 1997-05-22  
; NUMBER OF SEQ ID NOS: 94  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO: 27  
; LENGTH: 403  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: KS beta gene  
US-08-861-774E-27

Query Match 4.3%; Score 7; DB 4; Length 403;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 78 EHLPGRL 84  
|||||  
DB 59 EHLPGRL 65

RESULT 10

US-08-861-774E-25  
; Sequence 25, Application US/08861774E  
; Patent No. 6297007  
; GENERAL INFORMATION:  
; APPLICANT: Waters, Barbara  
; APPLICANT: Miao, Vivian

APPLICANT: Ho, Yap  
APPLICANT: Tong, Seow  
TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR  
TITLE OF INVENTION: BIOACTIVE MOLECULES  
FILE REFERENCE: 9993-006  
CURRENT APPLICATION NUMBER: US/08/861,774E  
CURRENT FILING DATE: 1997-05-22  
NUMBER OF SEQ ID NOS: 94  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 25  
LENGTH: 406  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: KS beta gene  
US-08-861-774E-25

Query Match 4.3%; Score 7; DB 4; Length 406;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 78 EHLPGRL 84  
|||||  
Db 60 EHLPGRL 66

RESULT 11  
US-08-861-774E-28  
Sequence 28, Application US/08861774E  
Patent No. 6297007  
GENERAL INFORMATION:  
APPLICANT: Waters, Barbara  
APPLICANT: Miao, Vivian  
APPLICANT: Ho, Yap  
TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR  
TITLE OF INVENTION: BIOACTIVE MOLECULES  
FILE REFERENCE: 9993-006  
CURRENT APPLICATION NUMBER: US/08/861,774E  
CURRENT FILING DATE: 1997-05-22  
NUMBER OF SEQ ID NOS: 94  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 407  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: KS beta gene  
US-08-861-774E-28

Query Match 4.3%; Score 7; DB 4; Length 407;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 78 EHLPGRL 84  
|||||  
Db 60 EHLPGRL 66

RESULT 12  
US-08-578-709-15  
Sequence 15, Application US/08578709  
Patent No. 5814509  
GENERAL INFORMATION:  
APPLICANT: TANABE, Tadashi  
TITLE OF INVENTION: PROSTACYCLIN SYNTHASE DERIVED FROM HUMAN  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington

STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/578,709  
FILING DATE: 28-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/00838  
FILING DATE: 27-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 114316/1994  
FILING DATE: 28-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gubinsky, Louis  
REGISTRATION NUMBER: 24,835  
REFERENCE/DOCKET NUMBER: Q40439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 500 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-578-709-15

Query Match 4.3%; Score 7; DB 2; Length 500;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVSQTTT 8  
|||||  
Db 323 PVSQTTT 329

RESULT 13  
US-09-238-448-1  
Sequence 1, Application US/09238448  
Patent No. 6238925  
GENERAL INFORMATION:  
APPLICANT: Sampson, Hugh A.  
TITLE OF INVENTION: Method for Determining Likelihood of Developing  
FILE REFERENCE: HS 100  
CURRENT APPLICATION NUMBER: US/09/238,448  
CURRENT FILING DATE: 1999-01-28  
EARLIER APPLICATION NUMBER: 60/073,171  
EARLIER FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-09-238-448-1

Query Match 3.7%; Score 6; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 CSRFPN 64  
|||||

Db 5 CSRFPPN 10

RESULT 14

US-09-238-448-5  
Sequence 5, Application US/09238448

Patent No. 6238923

GENERAL INFORMATION:

APPLICANT: Sampson, Hugh A.  
TITLE OF INVENTION: Method for Determining Likelihood of Developing  
TITLE OF INVENTION: Immunological Tolerance

FILE REFERENCE: HS 100

CURRENT APPLICATION NUMBER: US/09/238,448

EARLIER FILING DATE: 1999-01-28

EARLIER APPLICATION NUMBER: 60/073,171

NUMBER OF SEQ ID NOS: 10

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 5

LENGTH: 10

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: peptide

US-09-238-448-5

Query Match

Best Local Similarity 3.7%; Score 6; DB 4; Length 10;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CSRFPPN 64

Db 3 CSRFPPN 8

RESULT 15

US-07-876-941A-36

Sequence 36, Application US/07876941A

Patent No. 5885768

GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.

APPLICANT: Bradley, Daniel W.

APPLICANT: Tam, Albert W.

APPLICANT: Mitchell, Carl

TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and

TITLE OF INVENTION: Antibodies

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger &amp; Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/876,941A

FILING DATE: 01-MAY-1992

CLASSIFICATION: 435

APPLICATION NUMBER: US 822,335

FILING DATE: 17-JAN-1992

APPLICATION NUMBER: US 505,888

FILING DATE: 05-APRIL-1990

APPLICATION NUMBER: US 420,921

FILING DATE: 13-OCTOBER-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 367,486

FILING DATE: 16-JUNE-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 336,672

FILING DATE: 11-APRIL-1989

APPLICATION NUMBER: US 208,997

FILING DATE: 17-JUNE-1988

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 4600-0093.33

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Fig. 11, ORF 2, aa 341-354

US-07-876-941A-36

Query Match

Best Local Similarity 3.7%; Score 6; DB 2; Length 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTTAAT 11

Db 2 TTTAAT 7

Search completed: October 9, 2002, 01:47:15  
Job time: 11727 sec